

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 28, 2003, 15:27:11 ; Search time 31 Seconds
(without alignments)
1180.609 Million cell updates/sec

Title: US-10-055-364-24
Perfect score: 4547
Sequence: 1 MAGSLKRGSLVLAALWLYQV.....KRSQYSLNKAESVEMLNTL 865

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents, AA.*
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pdp.*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pdp.*
3: /cgn2_6/ptodata/2/iaa/6A_COMB.pdp.*
4: /cgn2_6/ptodata/2/iaa/6B_COMB.pdp.*
5: /cgn2_6/ptodata/2/iaa/PTCUS_COMB.pdp.*
6: /cgn2_6/ptodata/2/iaa/backfiles.pdp.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	4547	100.0	865	4	US-09-612-204B-24
2	1953.5	43.0	874	3	US-08-804-439A-16
3	1953.5	43.0	874	3	US-08-720-229-16
4	1949.5	42.9	808	3	US-08-804-439A-14
5	1949.5	42.9	808	3	US-08-720-229-14
6	1937	42.6	845	3	US-08-804-439A-94
7	1937	42.6	845	3	US-08-720-229-94
8	1848.5	40.7	857	1	US-08-220-151-10
9	1848.5	40.7	857	1	US-08-413-118-10
10	1848.5	40.7	857	3	US-08-804-439A-18
11	1848.5	40.7	857	3	US-08-360-107A-113
12	1848.5	40.7	857	3	US-08-473-446-10
13	1848.5	40.7	857	3	US-08-720-229-18
14	1839	40.4	856	3	US-08-486-099-103
15	1839	40.4	856	3	US-08-484-223B-103
16	1839	40.4	856	3	US-08-919-597-103
17	1839	40.4	856	3	US-08-475-668A-103
18	1839	40.4	856	3	US-08-485-551A-103
19	1839	40.4	856	3	US-08-471-913A-103
20	1839	40.4	856	3	US-08-485-264A-103
21	1839	40.4	856	4	US-08-474-349A-103
22	1839	40.4	856	4	US-08-470-896-103
23	1839	40.4	856	4	US-08-485-546A-103
24	1792	39.4	849	3	US-08-804-439A-17
25	1792	39.4	849	3	US-08-720-229-17
26	1787	39.3	874	3	US-08-804-439A-15
27	1787	39.3	874	3	US-08-720-229-15

28	1128	24.8	907	1	US-08-349-006-2	Sequence 2, Appl
29	1128	24.8	907	3	US-08-804-439A-19	Sequence 19, Appl
30	1128	24.8	907	3	US-08-720-229-19	Sequence 19, Appl
31	1128	24.8	907	4	US-09-171-699-2	Sequence 2, Appl
32	1128	24.8	907	5	PCT-US94-04180-2	Sequence 2, Appl
33	1121.5	24.7	906	1	US-08-220-151-9	Sequence 9, Appl
34	1121.5	24.7	906	1	US-08-413-118-9	Sequence 9, Appl
35	1121.5	24.7	906	3	US-08-473-446-9	Sequence 9, Appl
36	1050.5	23.1	830	3	US-08-804-439A-20	Sequence 20, Appl
37	1050.5	23.1	830	3	US-08-720-229-20	Sequence 20, Appl
38	1002	22.0	195	4	US-09-612-204B-2	Sequence 2, Appl
39	901	19.8	879	1	US-08-220-151-2	Sequence 2, Appl
40	901	19.8	879	1	US-08-220-151-3	Sequence 3, Appl
41	901	19.8	879	1	US-08-413-118-3	Sequence 2, Appl
42	901	19.8	879	1	US-08-413-118-3	Sequence 3, Appl
43	901	19.8	879	1	US-08-413-118-106	Sequence 106, App
44	901	19.8	879	3	US-08-473-446-2	Sequence 2, Appl
45	901	19.8	879	3	US-08-473-446-3	Sequence 3, Appl

ALIGNMENTS

RESULT 1
US-09-612-204B-24
; Sequence 24, Application US/09612204B
; Patent No. 6461811
; GENERAL INFORMATION:
; APPLICANT: Patience, Clive
; TITLE OF INVENTION: Swine Gamma Herpesvirus DNA and Methods of Use
; FILE REFERENCE: 61750-299
; CURRENT APPLICATION NUMBER: US/09/612.204B
; CURRENT FILING DATE: 2001-08-13
; PRIOR APPLICATION NUMBER: U.S. 60/142,736
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: U.S. 60/168,532
; PRIOR FILING DATE: 1999-12-02
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 24
; LENGTH: 865
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:Deduced amino acid sequence of porcine gamma herpesvirus gpb
; OTHER INFORMATION: gene
; OTHER INFORMATION: gene
US-09-612-204B-24

Query Match	100.0%;	Score 4547;	DB 4;	Length 865;
Best Local Similarity	100.0%;	Pred. No. 0;	Mismatches 0;	Indels 0;
Matches	865;	Conservative 0;		Gaps 0;
Qy	1	MAGSLKRGSLVLAALWLYQV	1	MAGSLKRGSLVLAALWLYQV
Db	1	MAGSLKRGSLVLAALWLYQV	1	MAGSLKRGSLVLAALWLYQV
Qy	61	RGNNENRDSBEQNKNIYGS	61	RGNNENRDSBEQNKNIYGS
Db	61	RGNNENRDSBEQNKNIYGS	61	RGNNENRDSBEQNKNIYGS
Qy	121	YKONIIPFMRVRKRVKVT	121	YKONIIPFMRVRKRVKVT
Db	121	YKONIIPFMRVRKRVKVT	121	YKONIIPFMRVRKRVKVT
Qy	181	LRNTGNNLLTYVDRDDIN	181	LRNTGNNLLTYVDRDDIN
Db	181	LRNTGNNLLTYVDRDDIN	181	LRNTGNNLLTYVDRDDIN
Qy	241	NCELMDFARSNPPFFV	241	NCELMDFARSNPPFFV
Db	241	NCELMDFARSNPPFFV	241	NCELMDFARSNPPFFV

```

QY 301 QNGRTVPLGKTRIFLDREYYTLTWSKHLKNMSYCPITLWKAFYNGIQTEHSGSYHFVAND 360
Db 301 QNGRTVPLGKTRIFLDREYYTLTWSKHLKNMSYCPITLWKAFYNGIQTEHSGSYHFVAND 360
QY 361 ITASFTTSKEDMKFNTTYHCLNEEKAEIEKKYAKVNSTHSGDLKYFKTDGGLYLW 420
Db 361 ITASFTTSKEDMKFNTTYHCLNEEKAEIEKKYAKVNSTHSGDLKYFKTDGGLYLW 420
QY 421 OPLIQNRLLDAKNNLNTYRSRROAESTTTPMMEMTNGAGGYSSENSITVAQVQY 480
Db 421 OPLIQNRLLDAKNNLNTYRSRROAESTTTPMMEMTNGAGGYSSENSITVAQVQY 480
QY 481 AYDNLRIINNILEDLSKAWCRQHRALVWNLSPKINPTSVMSIYNRRVSAKRIGDVI 540
Db 481 AYDNLRIINNILEDLSKAWCRQHRALVWNLSPKINPTSVMSIYNRRVSAKRIGDVI 540
QY 541 SVSNCIYVDQTSVSLHSLRLSASDEKCFSRPPVTFKFMNDSTIYKQGLGVNNEILLTT 600
Db 541 SVSNCIYVDQTSVSLHSLRLSASDEKCFSRPPVTFKFMNDSTIYKQGLGVNNEILLTT 600
QY 601 TYLETCOENTYFQAKTDMYIKYNYEHLKTVPLSSITTLDTFIALNFTLLENVDFKVE 660
Db 601 TYLETCOENTYFQAKTDMYIKYNYEHLKTVPLSSITTLDTFIALNFTLLENVDFKVE 660
QY 661 LYTRDEKRLSNVDFIETMFREYNVYAOVSGLRKLDDLTNRNQFVDAFGSLMDDLGA 720
Db 661 LYTRDEKRLSNVDFIETMFREYNVYAOVSGLRKLDDLTNRNQFVDAFGSLMDDLGA 720
QY 721 GQTVNNAVSGVATLFSISIVTGFINFKNPFGGMLMIIWIGVLFAYIYFLTKTKIYETAP 780
Db 721 GQTVNNAVSGVATLFSISIVTGFINFKNPFGGMLMIIWIGVLFAYIYFLTKTKIYETAP 780
QY 781 IKMIYPIDLKLREGKSEIAPISEEELERIVLAMHIHQONSHMETKTRKPKDSILTRA 840
Db 781 IKMIYPIDLKLREGKSEIAPISEEELERIVLAMHIHQONSHMETKTRKPKDSILTRA 840
QY 841 QNMLRKESGYSLNKNVSEVEMLNTL 865
Db 841 QNMLRKESGYSLNKNVSEVEMLNTL 865

```

RESULT 2

```

US-08-804-439A-16
; Sequence 16, Application US/08804439A
; Patent No. 6015565
; GENERAL INFORMATION:
; APPLICANT: Rose, Timothy M.
; APPLICANT: Bosch, Marnix L.
; APPLICANT: Strand, Kurt
; TITLE OF INVENTION: GLYCOPROTEIN B OF THE RPHV/KSHV
; TITLE OF INVENTION: SUBFAMILY OF HERPES VIRUSES
; NUMBER OF SEQUENCES: 113
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 4225 Executive Square, Ste 1400
; CITY: La Jolla
; STATE: CA
; COUNTRY: USA
; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/804.439A
; FILING DATE: February 21, 1997
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Haile, Lisa A.
; REGISTRATION NUMBER: 38,347
; REFERENCE/DOCKET NUMBER: 09176/004001
; TELECOMMUNICATION INFORMATION:

```

```

; TELEPHONE: (619) 678-5070
; TELEFAX: (619) 678-5099
; TELEX:
; INFORMATION FOR SEQ ID NO: 16:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 874 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-804-439A-16

```

```

Query Match 43.0%; Score 1953.5; DB 3; Length 874;
Best Local Similarity 46.1%; Pred. No. 2e-143;
Matches 403; Conservative 147; Mismatches 250; Indels 75; Gaps 18;

```

```

QY 11 VLALWLYQVALYSLSIAETGYTSPENTATWSTESPLTGHYGTGTHDSHGGRGNENRDS 70
Db 10 VLCLWCV--AALLCQGAQEVVA-----ETTTB---FATH-----RPEVVAE 46
QY 71 EQKNIIYGSPTFPYRVCASGV-GDVPRFQTDHVCPCDASDMVHSEGIILLIYKQNIIPFM 129
Db 47 ENPAMPF-----LPRFVCGASPTGGEIFRPLEESCPTEDKDHIEGIALIYKNTIYVPV 101
QY 130 FRVKRYKRVVTTSTVYNGIYSDSITNQHTFYKSIIEPWETEKMDTIYQCFNSLRNLNTG 189
Db 102 FNVKRYKRVVTTSTVYNGIYSDSITNQHTFYKSIIEPWETEKMDTIYQCFNSLRNLNTG 189
QY 190 LTYVDRDDINMTVLPQVDGVTDPVKRYGSGPELYLEPGWFGWYSYRRRTTVNCELMDFA 249
Db 162 NTYVDRDGNWETAFLKPADGLTSSITRYQSQPEVATPRNLLWSYTRTTVNCVETMSA 221
QY 250 RSNPPDFPVATGDTVENSFPWSGEDDHENKMKHEKWPVSVINNVKVDYQNRGTVPLG 309
Db 222 RSMKPEPFVTSVGDPTIENSFPFKENGTEPEKILKPHSIIQLLKNVAVTKY----GVGLG 277
QY 310 K---TRIFLDREYYTLTWSKHLKNMSYCPITLWKAFYNGIQTEHSGSYHFVANDITAS 365
Db 278 QADNATRFATFGDYSLSWKATENSYYCDLILWKGFSAIQTQHNSSLHFANDITAS 337
QY 366 TTSKEDMKFNTTYHCLNEEKAEIEKKYAKVNSTHSGDLKYFKTDGGLYLWQPLIQ 425
Db 338 STPLEBEANFNFTFKCIWNNTQBEIKLKEVEKTHRPNGTAKVYKTTGNLYIYWQPLIQ 397
QY 426 NRIIDAKNKLNN-----ETYSRSTRQAESTTTPMMEMTNGAGGYSSENS 472
Db 398 IDLLDTHAKLYNLNTATASPTSTPTSPRRRRDTS-----VSGGGNNGDNSTKEES 450
QY 473 ITVAQVOYAYDNLRIINNILEDLSKAWCRQHRALVWNLSPKINPTSVMSIYNRPVS 532
Db 451 VAASQVOYAYDNLRIINNILEDLSKAWCRQHRALVWNLSPKINPTSVMSIYNRPVS 510
QY 533 AKRIGDVISVNCIYVDQTSVSLHSLRLSASDEKCFSRPPVTFKFMNDSTIYKQGLV 592
Db 511 AKLIGDVSVSDCISVDQKSVFVHKMK-VPGKEDLCYTRPVGVGFKFNGSELFPAGQLP 569
QY 593 NNEILLTTTYLTCQENTYFQAKTDMYIKYNYEHLKTVPLSSITTLDTFIALNFTLLE 652
Db 570 RNEIVLSTSQVEYCOHSCHEFYQAGNQMYKYDYVYSTNLNLTDTPTLHTMITNLNLSVE 629
QY 653 NVDFKVIELYTRDEKRLSNVDFIETMFREYNVYAOVSGLRKLDDLTNRNQFVDAFG 711
Db 630 NIDFKVIELYKTEKRLSNVDFIETMFREYNVYAOVSGLRKLDDLTNRNQFVDAFG 689
QY 712 SLMDLLGAVGQTVNNAVSGVATLFSISIVTGFINFKNPFGGMLMIIWIGVLFAYIYFL 770
Db 690 DIMQDLGTIGKVVVNAVSGVSLFSGISVGSVIFPKNPFGGMLLIIWIGVVFVLEMT 749
QY 771 KTKIYETAPIKMIYPIDLKLREGKSEIAPISEEELERIVLAMHIHQONSHMETKTR 830
Db 750 RRSRIY-SAPIRMLYPGVERAAQEPGAH---PVSEDIQIRNLMGMHGFQORQRAEEARR 805
QY 831 DPK-----DSILTRAQNMLRKR---SGYSNLK 854

```



```
TELEFAX: (619) 678-5099
TELEX:
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 808 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-804-439A-14

Query Match      42.9%; Score 1949.5; DB 3; Length 808;
Best Local Similarity 46.6%; Pred. No. 3.6e-143;
Matches 373; Conservative 159; Mismatches 236; Indels 33; Gaps 10;

QY 71 EQNKNIYGSSTPYRVCASGVGVDFRQTDHVCPCDASDMVHSEGLIYKQNIIPFPMF 130
DB 28 EKNKTOAIYQYFYKRYVCSASTTGELFRDLDRTPCSTEDKVHKEGILLVYKKNIVPIYF 87
QY 131 RVRYKRVVTTSTVYNGIYSD--SITNQHTFYKSIETPWEKMDTIYQCFNSRLNTGN 188
DB 88 KVRKYKAITTSVRIFNGWTRREGVAITNKWELSRAPKYEIDIMDKTYQCHNCQIEVNGM 147
QY 189 LLTYVDRDDINMTVFLQPVGVTPDVKRYGSOPELYLEPGWFWGVSRRRTTVCNELMDMF 248
DB 148 LNSYDRDGNKNTVDLKPVDGUTGAIIRYISQPKVPADPGWLWGTTRTTTVCNEIVDMF 207
QY 249 ARSNPPDFPVFTATGDTVEMSPFSGEDDHNKMEKPMFVSVINNYKVVDYQNRGTVPL 308
DB 208 ARSADPYTYFTALGDTVEVSPFCVDNSCPNATDVL--VQIDLNHTVVDYGNRATSSQ 265
QY 309 GKTRIFLDREYTLNWEKHLKNMSYCPPLTWKAFYNGIOTHSYGVHVFANDITAGFTTS 368
DB 266 HKKRIFAHTLDYSVSWEAVNKSASVCSMVFWKSFQRAIQTEHDLTYHFIFANEITAGFSTV 325
QY 369 KEDMKFEINTTYHCLNEEIKAEIEKKYAKYNSTHSKYGDLKYFKTDGGLYLVWQPLIQNRL 428
DB 326 KEPLANFTSDYNCLMTHINTLEDKIARVNNHTPNGTAEYQTEGGMLLVWQPLIAIEL 385
QY 429 LDANKNLNNE-----TYSRRSRROAESTTDPMMEMTNGAGGEYSSENSITVAQVOYA 481
DB 386 EAMLEATTSPVTSAPTSSRSKRRAIRSIRDV-----SAG-----SENNVFLSQIOYA 434
QY 482 YDNLRIIRINILEDLSKAWCREQHRALVNELSKINPTSVMSIMYNRPVSAKRIGDVIS 541
DB 435 YDKLRQSIINNVLBELATWCREQVRQTVWVYEAIAKINPTSVMTAIGKPVSRKALGDVIS 494
QY 542 VSNICVVDQTSVSLHSLRLLSASDERCFRPPVTFKPMNDSTIYKGQLGVNNEILLTTT 601
DB 495 VTECINVDQSSVSIHSLK--TENNDICYSRPPVTFKFNSSQLFKQGLGARNEILLSSES 552
QY 602 YLETCQENTYYPQAKTDMYIKNYEHLKTVPLSSITLDTFALNPTLLENVDFKVEL 661
DB 553 LVENCHQNAETFTAKNETYHFKNYVHVELTPVNNISTLDTFLALNLTFTENIDFKAVEL 612
QY 662 YTRDEKSLNVFDETFRFNYNYAORVSLGRKDLDLDS--TNRNQPVDAFGSLMDDLGA 719
DB 613 YSSGERKANVFDLTFRFRNYNYAQISGLRKD--FNSQRNDRRIIQDFSEILADLGS 671
QY 720 VQGTVVNAVSGVATLFSISVTGFINFKNPFGGMLMIIVIGVLFYAIYFLTKTKIYETA 779
DB 672 IGKVINNVASGAFSLFGGIVTGILNFKNPLGGMFTFLGAVILLVILVRRTNNMSQA 731
QY 780 PIKMIYPEIDKLEREGKSIAPISEEELERIVLAMIHOONSHMETKTRKDPKDSILTR 839
DB 732 PIRMIYDPVFK-----SKSTVTWPEPETIKQIILGHMNMQOEAYKKKEQOARAPSFQR 786
QY 840 -AQNMLKRSYSNKLNAESV 859
DB 787 AATFLKRSGYKQISTEDKI 807
```

RESULT 5

```
US-08-720-229-14
Sequence 14, Application US/08720229
Patent No. 6022542
GENERAL INFORMATION:
APPLICANT: Rose, Timothy M.
APPLICANT: Bosch, Marnix L.
APPLICANT: Strand, Kurt
TITLE OF INVENTION: GLYCOPROTEIN B OF THE RHFV/KSHV
TITLE OF INVENTION: SUBFAMILY OF HERPES VIRUSES
NUMBER OF SEQUENCES: 100
CORRESPONDENCE ADDRESS:
ADDRESSEE: Morrison & Foerster
STREET: 755 Page Mill Road
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304-1018
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/720,229
FILING DATE: 26-SEP-1996
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Schiff, J. Michael
REGISTRATION NUMBER: 40,253
REFERENCE/DOCKET NUMBER: 29938-20002.00
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 813-5600
TELEFAX: (415) 494-0792
TELEX: 706141
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 808 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-720-229-14

Query Match      42.9%; Score 1949.5; DB 3; Length 808;
Best Local Similarity 46.6%; Pred. No. 3.6e-143;
Matches 373; Conservative 159; Mismatches 236; Indels 33; Gaps 10;

QY 71 EQNKNIYGSSTPYRVCASGVGVDFRQTDHVCPCDASDMVHSEGLIYKQNIIPFPMF 130
DB 28 EKNKTOAIYQYFYKRYVCSASTTGELFRDLDRTPCSTEDKVHKEGILLVYKKNIVPIYF 87
QY 131 RVRYKRVVTTSTVYNGIYSD--SITNQHTFYKSIETPWEKMDTIYQCFNSRLNTGN 188
DB 88 KVRKYKAITTSVRIFNGWTRREGVAITNKWELSRAPKYEIDIMDKTYQCHNCQIEVNGM 147
QY 189 LLTYVDRDDINMTVFLQPVGVTPDVKRYGSOPELYLEPGWFWGVSRRRTTVCNELMDMF 248
DB 148 LNSYDRDGNKNTVDLKPVDGUTGAIIRYISQPKVPADPGWLWGTTRTTTVCNEIVDMF 207
QY 249 ARSNPPDFPVFTATGDTVEMSPFSGEDDHNKMEKPMFVSVINNYKVVDYQNRGTVPL 308
DB 208 ARSADPYTYFTALGDTVEVSPFCVDNSCPNATDVL--VQIDLNHTVVDYGNRATSSQ 265
QY 309 GKTRIFLDREYTLNWEKHLKNMSYCPPLTWKAFYNGIOTHSYGVHVFANDITAGFTTS 368
DB 266 HKKRIFAHTLDYSVSWEAVNKSASVCSMVFWKSFQRAIQTEHDLTYHFIFANEITAGFSTV 325
QY 369 KEDMKFEINTTYHCLNEEIKAEIEKKYAKYNSTHSKYGDLKYFKTDGGLYLVWQPLIQNRL 428
DB 326 KEPLANFTSDYNCLMTHINTLEDKIARVNNHTPNGTAEYQTEGGMLLVWQPLIAIEL 385
QY 429 LDANKNLNNE-----TYSRRSRROAESTTDPMMEMTNGAGGEYSSENSITVAQVOYA 481
DB 386 EAMLEATTSPVTSAPTSSRSKRRAIRSIRDV-----SAG-----SENNVFLSQIOYA 434
QY 482 YDNLRIIRINILEDLSKAWCREQHRALVNELSKINPTSVMSIMYNRPVSAKRIGDVIS 541
DB 435 YDKLRQSIINNVLBELATWCREQVRQTVWVYEAIAKINPTSVMTAIGKPVSRKALGDVIS 494
QY 542 VSNICVVDQTSVSLHSLRLLSASDERCFRPPVTFKPMNDSTIYKGQLGVNNEILLTTT 601
DB 495 VTECINVDQSSVSIHSLK--TENNDICYSRPPVTFKFNSSQLFKQGLGARNEILLSSES 552
QY 602 YLETCQENTYYPQAKTDMYIKNYEHLKTVPLSSITLDTFALNPTLLENVDFKVEL 661
DB 553 LVENCHQNAETFTAKNETYHFKNYVHVELTPVNNISTLDTFLALNLTFTENIDFKAVEL 612
QY 662 YTRDEKSLNVFDETFRFNYNYAORVSLGRKDLDLDS--TNRNQPVDAFGSLMDDLGA 719
DB 613 YSSGERKANVFDLTFRFRNYNYAQISGLRKD--FNSQRNDRRIIQDFSEILADLGS 671
QY 720 VQGTVVNAVSGVATLFSISVTGFINFKNPFGGMLMIIVIGVLFYAIYFLTKTKIYETA 779
DB 672 IGKVINNVASGAFSLFGGIVTGILNFKNPLGGMFTFLGAVILLVILVRRTNNMSQA 731
QY 780 PIKMIYPEIDKLEREGKSIAPISEEELERIVLAMIHOONSHMETKTRKDPKDSILTR 839
DB 732 PIRMIYDPVFK-----SKSTVTWPEPETIKQIILGHMNMQOEAYKKKEQOARAPSFQR 786
QY 840 -AQNMLKRSYSNKLNAESV 859
DB 787 AATFLKRSGYKQISTEDKI 807
```

Db 386 BEAMLEATSPVTPSAPTSSRSKRAIRSDV-----SAG-----SENNVFLSQIQYA 434
Qy 482 YDNLRIRINILEDISKACWROHRAALVWNLKINPTSMVMYINRPVSAKRIGDVIS 541
Db 435 YDKLRQSNINVLLELAITWCROQVOTWVYIAKINPTSMVIAIGKPVSKAUGDVIS 494
Qy 542 VENCIVDQTSVSLHKSLLLSASDEKCFSPVTFKFMNDSTIYKQGLGVNNEILLTTT 601
Db 495 VTECINVDQSSVSIHSLK--TENNDICVSRPPVTFKFNSSQLFKGQLGARNEILLSES 552
Qy 602 YLETCOENETYYFOAKTDMYIYKNVEHLKTVPLSSITLDTFIALNFTLLENVDFKVEL 661
Db 553 LVENCHQNAETFTAKNETHYKPNVHVETLFPVNNISLDTFLALNLFIEINIDFKAVEL 612
Qy 662 YTRDEKRLSNVFDIETMPREYNVYQVRVSGRLKDLDS--NRRNQVDVAFGLMDDLGA 719
Db 613 YSSGERKANVDELTMFREYNVYQASISGLRKD-FDMSQRNRRRIIQDFSEILLADLOS 671
Qy 720 VQGTVVNAVSGVATLFFSSIVTGFINFKNPFGLMLMIIVIGLVFAIYFLTKTKIYETA 779
Db 672 IGKVIWVASGAFSLFGGIVTGILNFKNPLGLGMFTLLIGAVIILVILLVRRTNMMSQA 731
Qy 780 PIKMIYPEIDKLEREGKSEIAPISEEELERIVLAHMHQONSHMETKTRKDPKDSILTR 839
Db 732 PIRMIYDVEK-----SKSTVTMPPEPTIKOILLGHNMQOEAYKKKEEQAARPSIFRQ 786
Qy 840 -AQNMLRKRSGYSLKNAESV 859
Db 787 AAETFLRRSGYKQISTEDKI 807

RESULT 6

US-08-804-439A-94
; Sequence 94, Application US/08804439A
; Patent No. 6015565
; GENERAL INFORMATION:
; APPLICANT: Rose, Timothy M.
; APPLICANT: Bosch, Marnix L.
; APPLICANT: Strand, Kurt
; TITLE OF INVENTION: GLYCOPROTEIN B OF THE RPHV/KSHV
; TITLE OF INVENTION: SUBFAMILY OF HERPES VIRUSES
; NUMBER OF SEQUENCES: 113
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 4225 Executive Square, Ste 1400
; CITY: La Jolla
; STATE: CA
; COUNTRY: USA
; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/804,439A
; FILING DATE: February 21, 1997
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Haile, Lisa A.
; REGISTRATION NUMBER: 38,347
; REFERENCE/DOCKET NUMBER: 09176/004001
; TELEPHONE: (619) 678-5070
; TELEFAX: (619) 678-5099
; TELEX:
; INFORMATION FOR SEQ ID NO: 94:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 845 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; FEATURE:

; NAME/KEY: Modified-site
; LOCATION: 841
; OTHER INFORMATION: /note= "Proline or Leucine
; OTHER INFORMATION: depending on codon"
US-08-804-439A-94

Query Match 42.6%; Score 1937; DB 3; Length 845;
Best Local Similarity 48.5%; Pred. No. 3.7e-142;

Matches 382; Conservative 135; Mismatches 237; Indels 34; Gaps 10;

Qy 83 PPRVCSASGVDPVFRQTDHVCPSASDMVHSEGLLIYKONIIPFPRVRKYKRVVTS 142
Db 63 YQFRVCSASITGELFRNLEQTCPTDKVKYHQEGILLVYKKNIVPHIFKVRVKIATSV 122
Qy 143 TVYNGIYSDSITNOHTFYKSIPEWETEKMDTIYOCFNSLRNLNTGNTLLTYVDRODINMTV 202
Db 123 TVYGLTESAITNKEYELPRPVLYEISHMDSTYQCFSSMKVNVNGVENTFTFDRDDVNTTV 182
Qy 203 FLQPDVGVTDPVKRYGSOPELYLEPGFWGYSRRRTTVNCELMDMFARSNPPFFVVTAT 262
Db 183 FLQPEGLTDNIQRYFQPVYAEFGWFPGLYRVRTTVNCEIVDMIARSAEPYNYFVTS 242
Qy 263 GDTVEMGPFWSGDEDDHKNMHEKWPV-VSYNNYKVVVDYQNRGTVPGLKTRIFLDREY 321
Db 243 GDTVEVSPFCYNSSCSSTPSNKNGLSVQVNLNHTVTVYSDRGTSPQNRIFVETGAYT 302
Qy 322 LSWEKHLKNMSYCPCLTLWKAFYNGIOTHSYGHFVANDITASFTTSKEDMKENFTVHC 381
Db 303 LSWASESKTTAVCPALWKTFFRSIQTHEDSPHFVAEITATFTAPLTPVANFTDITYSC 362
Qy 382 LNEEIKAEIEKKYAKVNSTHSKYGDLKYFKTGDGLYVWQPLIQNRLLDAKNKLNNET-- 439
Db 363 LSTDINTLNAASKAKLASTHVPGNTVOYFHTTGLYLVWQPMASINLTHAGDSGNPTSS 422
Qy 440 -----YSRRRQAESTTDPMMWMTGNAGGEYSSENSITVAQVQYAYDNLRIR 488
Db 423 PPSASPMWTSASRRKERSASTA-----AAGGGSTDN-LSVTOLQFAYDKLRDG 471
Qy 489 INNILEDLSKAWCREQHRALVWNLKINPTSMVMYINRPVSAKRIGDVISVNCIIV 548
Db 472 INQVLEELSRACWCREQVRDNLMMYELSKINPTSMVTAIYGRPVSAKFVGDAISVTECIN 531
Qy 549 DQTSVSLHKSRLLSASDEKCFSPVTFKFMNDSTIYKQGLGVNNEILLTTTLETCE 608
Db 532 DQSSVNIHSLR--TNSKDVYARPLVTFKFLSSNLTGOLGARNIILLTNQVEICKD 589
Qy 609 NTEYFOAKTDMYIYKNVEHLKTVPLSSITLDTFIALNFTLLENVDFKVELYTRDEKR 668
Db 590 TCEHYFITRNETLVYKDYAVLRTINTDTLNTFIALNLSFIQNIIDFKALELYSSAEKR 649
Qy 669 L-SNVFDIETMPREYNVYQVRVSGRLKDLDS--LSTNRNQVDVAFGLMDDLGA VGTWVN 726
Db 650 LASSVFDLETMPREYNYYTHRLAGLREDLONTDMNKEFRVLDSEIVADLGGTGVN 709
Qy 727 AVSGVATLFFSSIVTGFINFKNPFGLMLMIIVIGLVFAIYFLTKTKIYETAPIKMIY 786
Db 710 VASSVTLGSLVTFGFINFKPLGGLMLIIVIAILLFMSLRRTNTIAQAPVKMIY 769
Qy 787 EIDKLKEREGKSEIAPISEEELERIVLAHMHQONSHMETKTRKDPKDSILTR AQNMLRK 846
Db 770 DVDR--RAPPSSGAP-TREEIKNILLGMHQLQOEERQKADDLKSKTSPSFQRTANGLRQ 825
Qy 847 R-SGYSNL 853
Db 826 RLGRYKPL 833

RESULT 7

US-08-720-229-94
; Sequence 94, Application US/08720229
; Patent No. 6022542
; GENERAL INFORMATION:
; APPLICANT: Rose, Timothy M.


```
Query Match      40.7%; Score 1848.5; DB 1; Length 857;
Best Local Similarity 46.1%; Pred. No. 2.9e-135;
Matches 368; Conservative 148; Mismatches 241; Indels 41; Gaps 12;

Qy 81 STFPYRVCSAGVGDFRQTDHVCDD-ASDMVHSEGLLIYKQNIIPFMRVRYKRVV 139
Db 44 TSFPRVCELSSHGLDFRSDIQCPSTRENHTEGLLMVFKDNIIPYSFKVRSYTKIV 103
Qy 140 TTSTVYNGIYSDSTNQHTEFYKSIPEWETEKMDTIYOCFNSRLNTGNNLLTYVDRDDIN 199
Db 104 TNILYNGWADSVNRRHEKFSVDSYETQMDTIYOCYNVMTKDGLTTRYVDRDGVN 163
Qy 200 MTFVLQPDVGVTPDKRYGQPELYLEPGWFGSYRRRTTNCLEMDMFARSNPPDFV 259
Db 164 ITVNLKPTGGLANGVRYASQTELYDAPGWLITWYRTTIVNCLITDMAKSNPFDFV 223
Qy 260 TATGTVEMSPFMSGEDDHENKMKHPFVSVINNYKVVDYQNRGTVPGLKTRIFLDREE 319
Db 224 TTTGQTVEMSPFYDGK--NKETFERADSFHVRTNYKIVDYNRGTNPQGERAFDOKGT 281
Qy 320 YTLSEKHLKN-MSYCPILTAKAFYNGIOTESHGSHYFVANDITASFTTSKEDMKFNTT 378
Db 282 YTLSEKHLKN-MSYCPILTAKAFYNGIOTESHGSHYFVANDITASFTTSKEDMKFNTT 378
Qy 379 YHCLNEEIKAEIEKKYAKVNSTHSGYGD-LKYFKTDGGLYLVWQPLIQNRLDKN--KL 435
Db 340 FKCIIEQVNTMKHEKYEAQVDRYTKQEALTYFITSGLLLAWMLPLTPRSLATVKNLTEL 399
Qy 436 NNETY-----RRSRQAESTTDPMMEMTONGAGGESSNSI 473
Db 400 TPTSPSPSPSPAPSAARGSTPAALVLRRRRRDAGNATP---VPPTAPCKSLGLTNNP 456
Qy 474 TVAQVOYAYDNLIRINNILEDLSKAWCREQRAALVWNLKINPTSVMSIYNRPVSA 533
Db 457 ATVQIQFAYDSLRRQINRMGLDLARAWCLEQKQNMVRLTKINPTTVMSSYIGKAVAA 516
Qy 534 KRIGDVISVNCIVDQTSVSLHSLRLLSASDEKCFSPRPVTPKFMNDSTIYKGLGVN 593
Db 517 KRLGDVISVOCVPVQATVTLRSMR--VPGSETWCYSRPLVSFINDTITYEGQUGTD 575
Qy 594 NEILLTITTYLETCQNTYFYQAKTDMYIKYNEHLKTVPLSSITLDTFIALNFTLLEN 653
Db 576 NEIFLTKWTEVCQATSYFQSGNEIHVNDYHHFKTIELDGIATLQTFISLNTSLIEN 635
Qy 654 VDFKVIELYTDEKLSNVFIETMFRYNYAQRVSLGRKOLL--LSTNRNQFVDAFGS 712
Db 636 IDFASLELYSRDEQASNVDFLEGIFREYNFQAQNIAGLRKLDLONAVSNGRNQFVGLGE 695
Qy 713 LMDDLGAQGVTVNAVSGVATLFFSSIVTGFINFKNPFGLMIIIVGLFAIYFVLTCK 772
Db 696 LMDSLGSGVQITNLVSTVGLFSSLSVSGFISFFKNPFGGLMILVAVAGVILVISLRR 755
Qy 773 TKIVETAPIKMYPIDIKLEREGKSE---TAPISEEELERIVLAMHIQOONSHMETKTR 829
Db 756 TRQMSQOQVOMLYPGIDELAQHASGEGGINPISKTELQAIMLA--LHEQNOQKRAAQ 813
Qy 830 KDPKDSILTRAQWMLRKR 847
Db 814 RAAGPSVASRALQARDR 831
```

RESULT 9

```
US-08-413-118-10
; Sequence 10, Application US/08413118
; Patent No. 5688920
; GENERAL INFORMATION:
; APPLICANT: PAOLETTI, ENZO
; APPLICANT: LIMBACH, KEITH J.
; TITLE OF INVENTION: NUCLEOTIDE AND AMINO ACID SEQUENCES OF
; CANINE HERPESVIRUS gb, gc, AND gd AND USES THEREFOR
; NUMBER OF SEQUENCES: 128
; CORRESPONDENCE ADDRESS:
```

```
; ADDRESSEE: CURTIS, MORRIS & SAFFORD, P.C.
; STREET: 530 FIFTH AVENUE, 25TH FLOOR
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: UNITED STATES OF AMERICA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/413,118
; FILING DATE: 29-MAR-1995
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/220,151
; FILING DATE: 30-MAR-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: FROMMER, WILLIAM S.
; REGISTRATION NUMBER: 25,506
; REFERENCE/DOCKET NUMBER: 454310-2670
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 840-3333
; TELEFAX: (212) 840-0712
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 857 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FRAGMENT TYPE: N-terminal
; US-08-413-118-10
```

```
Query Match      40.7%; Score 1848.5; DB 1; Length 857;
Best Local Similarity 46.1%; Pred. No. 2.9e-135;
Matches 368; Conservative 148; Mismatches 241; Indels 41; Gaps 12;

Qy 81 STFPYRVCSAGVGDFRQTDHVCDD-ASDMVHSEGLLIYKQNIIPFMRVRYKRVV 139
Db 44 TSFPRVCELSSHGLDFRSDIQCPSTRENHTEGLLMVFKDNIIPYSFKVRSYTKIV 103
Qy 140 TTSTVYNGIYSDSTNQHTEFYKSIPEWETEKMDTIYOCFNSRLNTGNNLLTYVDRDDIN 199
Db 104 TNILYNGWADSVNRRHEKFSVDSYETQMDTIYOCYNVMTKDGLTTRYVDRDGVN 163
Qy 200 MTFVLQPDVGVTPDKRYGQPELYLEPGWFGSYRRRTTNCLEMDMFARSNPPDFV 259
Db 164 ITVNLKPTGGLANGVRYASQTELYDAPGWLITWYRTTIVNCLITDMAKSNPFDFV 223
Qy 260 TATGTVEMSPFMSGEDDHENKMKHPFVSVINNYKVVDYQNRGTVPGLKTRIFLDREE 319
Db 224 TTTGQTVEMSPFYDGK--NKETFERADSFHVRTNYKIVDYNRGTNPQGERAFDOKGT 281
Qy 320 YTLSEKHLKN-MSYCPILTAKAFYNGIOTESHGSHYFVANDITASFTTSKEDMKFNTT 378
Db 282 YTLSEKHLKN-MSYCPILTAKAFYNGIOTESHGSHYFVANDITASFTTSKEDMKFNTT 378
Qy 379 YHCLNEEIKAEIEKKYAKVNSTHSGYGD-LKYFKTDGGLYLVWQPLIQNRLDKN--KL 435
Db 340 FKCIIEQVNTMKHEKYEAQVDRYTKQEALTYFITSGLLLAWMLPLTPRSLATVKNLTEL 399
Qy 436 NNETY-----RRSRQAESTTDPMMEMTONGAGGESSNSI 473
Db 400 TPTSPSPSPSPAPSAARGSTPAALVLRRRRRDAGNATP---VPPTAPCKSLGLTNNP 456
Qy 474 TVAQVOYAYDNLIRINNILEDLSKAWCREQRAALVWNLKINPTSVMSIYNRPVSA 533
Db 457 ATVQIQFAYDSLRRQINRMGLDLARAWCLEQKQNMVRLTKINPTTVMSSYIGKAVAA 516
Qy 534 KRIGDVISVNCIVDQTSVSLHSLRLLSASDEKCFSPRPVTPKFMNDSTIYKGLGVN 593
```


Db	576	NEIFLTKMTEVCQATSYQSFQSGNEIHVYNDYHHFKTIEDGIATQTFISLNTSLIEN	635
Qy	654	VDKVIETLYTRDEKRLSNVFDIETMREYNYAORVSGRLKDLDD-LSTNRNQFVDAFGS	712
Db	636	IDFASLELYSRDEQASNVFDELFREYFNOAQNIAGRLKDLNVAENGRNQFVDGLGE	695
Qy	713	LMDDLGAQGQTVVNAVSGVATLFSISVTGFINFKNPFGGMLMIVVIGVLFAYFLTKK	772
Db	696	LMDSLGSVQSITNLVSTVGLFSSLVSGFISFFKNPFGGMLILVLVAGVVILVSLTRR	755
Qy	773	TKIYETAPIKMYPEIDKLEREGKSE---IAPISEEELERIVLAMIHQONSHMETKTR	829
Db	756	TRMSQQPVMQLYPGIDELAQOCHASGEGPINPISKTELQAIMLA--LHEQNOBQKRAAQ	813
Qy	830	KDPKDSILTRAQNMLRKR	847
Db	814	RAAGFSVASRALQAARDR	831
RESULT 12			
US-08-473-446-10			
: Sequence 10, Application US/08473446			
: Patent No. 6017542			
: GENERAL INFORMATION:			
: APPLICANT: PAOLETTI, ENZO			
: APPLICANT: LIMBACH, KEITH J.			
: TITLE OF INVENTION: NUCLEOTIDE AND AMINO ACID SEQUENCES OF			
: TITLE OF INVENTION: CANINE HERPESVIRUS GB, GC, AND GD AND USES THEREFOR			
: NUMBER OF SEQUENCES: 128			
: CORRESPONDENCE ADDRESS:			
: ADDRESSES: CURTIS, MORRIS & SAFFORD, P.C.			
: STREET: 530 FIFTH AVENUE, 25TH FLOOR			
: CITY: NEW YORK			
: STATE: NEW YORK			
: COUNTRY: UNITED STATES OF AMERICA			
: ZIP: 10036			
: COMPUTER READABLE FORM:			
: MEDIUM TYPE: Floppy disk			
: COMPUTER: IBM PC compatible			
: OPERATING SYSTEM: PC-DOS/MS-DOS			
: SOFTWARE: Patent in Release #1.0, Version #1.30			
: CURRENT APPLICATION DATA:			
: APPLICATION NUMBER: US/08/473.446			
: FILING DATE:			
: CLASSIFICATION:			
: PRIORITY APPLICATION DATA:			
: APPLICATION NUMBER: 08/413,118			
: FILING DATE:			
: ATTORNEY/AGENT INFORMATION:			
: NAME: FROMMER, WILLIAM S.			
: REGISTRATION NUMBER: 25,506			
: REFERENCE/DOCKET NUMBER: 454310-2670			
: TELECOMMUNICATION INFORMATION:			
: TELEPHONE: (212) 840-3333			
: TELEFAX: (212) 840-0712			
: INFORMATION FOR SEQ ID NO: 10:			
: SEQUENCE CHARACTERISTICS:			
: LENGTH: 857 amino acids			
: TYPE: amino acid			
: STRANDEDNESS: single			
: TOPOLOGY: linear			
: MOLECULE TYPE: peptide			
: FRAGMENT TYPE: N-terminal			
US-08-473-446-10			
Query Match 40.7%; Score 1848.5; DB 3; Length 857;			
Best Local Similarity 46.1%; Pred. No. 2.9e-135;			
Matches 368; Conservative 148; Mismatches 241; Indels 41; Gaps 12;			
Qy	81	STPPYVCASGVDVFRQTHVCPD-ASDMVHSEGLLTYKQNIIPFMRVRYKRVV	139
Db	44	TSFPFVCLSHGDLFRSSDIQCPSTRENHTEGLLMVFKONIIPYSFKVRSYTKIV	103
Qy	140	TTSTVNGIYSDITNOHTFYKSIPEWTEKMDTIYQCFNSLRNLNTGNNLLTYVDRDDIN	199
Db	104	TNILLINGHYASVTRNHEKFSVSYETDQMDTIYQCNVAVKTKOGLTRVYVDRGVN	163
Qy	200	MTVLFQVDPVTPDKRYQSPELYLEPGWFSGSYRRRTTNCVELMDMFARNSPPDFV	259
Db	164	ITVNLKPTGLANGVRRYASQELYDAPGWLITVTRITVTVNCLITDMWAKSNSPPDFV	223
Qy	260	TATGDTVENSPPWSGEDDHENKMKHPFVSVINNYKVVDVQNRGTVPGLKTRIFLDREE	319
Db	224	TTTGQTVEMSPYDGR--NKETFHERADSFHVRTNYKIVDYVDRNRTNPGQERRAFLDKGT	281
Qy	320	YTLWSKHLKN--MSYCPLTLKAFYNGLOTEHSGSYHVFANDITASFTTSKEDMKENYTT	378
Db	282	YTLWSK--LENRTAYCPLQHWQTFDSTATETGKSIHFVTEDEGTSFVNTTVGIELPDA	339
Qy	379	YHCLNEEIKAEIKKYAKVNSTHYSKYGD--LKVFKTDDGGLYLVWQPLIQNRLLDKN--KL	435
Db	340	FKCIEQVKNTHKYEAVQDRTYTKGEAITVITSGLLLAWLPLTPRSLATYKNLTEL	399
Qy	436	NNETYS-----RRSRQEAESTTDDMMWTGNAGGEVSSNSI	4

```
Qy 140 TTSTVYNGIYSDITNOHTFYKSEPWETEKMDTIYOCFNSLRNLTNGNLLTVVDRDDIN 199
Db 104 TNILYNGWADSVTNRHEKEFSVDSYETQMDTIYOCYNNAVYTKDGLTRVYVDRDGVN 163
Qy 200 MTVFLQPDVGVTPDKRYGQPELYLEPGFWGYSRRRTTVNCELMDMFARSNPPDFV 259
Db 164 ITVNLKFTGGLANGVRVYASQTELYDAPGWLWYTRTITVNCILITDMMAKSNPPDFV 223
Qy 260 TATGDTVEMSPFSGEDDHENKMKHEKFWFVSINNYKVVDYQNRGTVPGLKTRIFLDREE 319
Db 224 TTTGQTVEMSPFVDGK--NKETFERADSFHVRTNYKIVDYDNRGTNPQGERAFLDKGT 281
Qy 320 YTLSEWHLKN-MSYCPCLTLWKAFYNGIOTESHSGSYHFVANDITASFTTSKEDMKENNT 378
Db 282 YTLSEWHLKN-MSYCPCLTLWKAFYNGIOTESHSGSYHFVANDITASFTTSKEDMKENNT 378
Qy 379 YHCLNEEIKAEIKYAKVNSTHSGYD-LKYFKTDGGLYLVNQPLIQNRLLDKN--KL 435
Db 340 FKCIIEQVNTMKHEKFAVQDRYTKGQEAITYFITSGGLLLAWLPLTPRSLATVKNLTTEL 399
Qy 436 NNETY-----RRSRQAESTTDPMMEMTNGAGGEYSSENSI 473
Db 400 TPTSPSPSPSPAPSAARGSTPAALVRRRRRDAGNATTP---VPPTAPGKSLGLTLNPP 456
Qy 474 TVAQVAYDNLRIIRNNILEDLSKAWCREQHRALVWNLKSKINPTSVMSMYNRPVSA 533
Db 457 ATVQIOFAYDNLRIIRNNILEDLSKAWCREQHRALVWNLKSKINPTSVMSMYNRPVSA 536
Qy 534 KRIGDVISVNCIVVDQTSVSLHSLRLSASDEKCFSPRPVTPFKMNDSTIYKQGLGVN 593
Db 517 KRLGDVISVOCVPVQATVTLRKSMT--VPGSETMTCYSRPLVSFNFINDTKYEGQLGTD 575
Qy 594 NEILLTTLTLETCOENTYFYFOAKTDMYIKYKNEHLKTVPLSSITLTDITFIALNFTLEN 653
Db 576 NEIFLTKMTKEVCQATSYFYQSGNEIHVYNDVHFKETIELDGIATLQTFISLNTSLIEN 635
Qy 654 VDPKVIETRDEKRLGNVFDIETMFREYNVYQVRVGLSKRLDLD--LSTNRNOFVDAFGS 712
Db 636 IDFASLEYSRDEQASNVFDLEGIHFREYNFQAGNIAGLRKLDLNAVSGNRQFVQDLGE 695
Qy 713 LMDDLGAVGQTVVNAVSGVATLFSSIVTGFINKPFGGMLMIIVVIGVLFALYFLTKK 772
Db 696 LMDSLGSGVQSIITNLVSTVGLFSLVSGFISFPKPFPGMLIILVAVGVILVISLTR 755
Qy 773 TKIYETAPIKWIPEIDKLKEREKSE---IAPISEEELERIVLAMIHOONSHMETKR 829
Db 756 TRQMSQOPVQMLYPCIGIDELAQHASGEGPGINPISKTELQAIMLA--LHEQNOEQKRAAQ 813
Qy 830 KDPKDSILTQAOQNLRR 847
Db 814 RAAGPSVASRALQAARD 831
```

RESULT 13

US-08-720-229-18

Sequence 18, Application US/08720229

Patent No. 6022542

GENERAL INFORMATION:

APPLICANT: Rose, Timothy M.

APPLICANT: Bosch, Marix L.

APPLICANT: Strand, Kurt

TITLE OF INVENTION: GLYCOPROTEIN B OF THE RHV/KSHV

SUBFAMILY OF HERPES VIRUSES

NUMBER OF SEQUENCES: 100

CORRESPONDENCE ADDRESS:

ADDRESS: Morrison & Foerster

CITY: Palo Alto

STATE: CA

COUNTRY: USA

ZIP: 94304-1018

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

```
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/720,229
FILING DATE: 26-SEP-1996
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Schiff, J. Michael
REGISTRATION NUMBER: 40,253
REFERENCE/DOCKET NUMBER: 29938-20002.00
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 813-5600
TELEFAX: (415) 494-0792
TELEX: 706141
INFORMATION FOR SEQ ID NO: 18:
SEQUENCE CHARACTERISTICS:
LENGTH: 857 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-720-229-18
```

Query Match 40.7%; Score 1848.5; DB 3; Length 857;

Best Local Similarity 46.1%; Pred. No. 2.9e-135; Mismatches 241; Indels 41; Gaps 12;

Matches 368; Conservative 148; Mismatches 241; Indels 41; Gaps 12;

Qy 81 STPPYVVCASGVGVFRFQTDHVCDD-ASDMVHSEGIILYKQNIIPFMFRVRYKRVV 139

Db 44 TSPPFVVCVCLSHGDLFRSSDIQCPSPCTRENHTSGLLVKFNIIYPSFKVRSYTKIV 103

Qy 140 TTSTVYNGIYSDITNOHTFYKSEPWETEKMDTIYOCFNSLRNLTNGNLLTVVDRDDIN 199

Db 104 TNILYNGWADSVTNRHEKEFSVDSYETQMDTIYOCYNNAVYTKDGLTRVYVDRDGVN 163

Qy 200 MTVFLQPDVGVTPDKRYGQPELYLEPGFWGYSRRRTTVNCELMDMFARSNPPDFV 259

Db 164 ITVNLKFTGGLANGVRVYASQTELYDAPGWLWYTRTITVNCILITDMMAKSNPPDFV 223

Qy 260 TATGDTVEMSPFSGEDDHENKMKHEKFWFVSINNYKVVDYQNRGTVPGLKTRIFLDREE 319

Db 224 TTTGQTVEMSPFVDGK--NKETFERADSFHVRTNYKIVDYDNRGTNPQGERAFLDKGT 281

Qy 320 YTLSEWHLKN-MSYCPCLTLWKAFYNGIOTESHSGSYHFVANDITASFTTSKEDMKENNT 378

Db 282 YTLSEWHLKN-MSYCPCLTLWKAFYNGIOTESHSGSYHFVANDITASFTTSKEDMKENNT 378

Qy 379 YHCLNEEIKAEIKYAKVNSTHSGYD-LKYFKTDGGLYLVNQPLIQNRLLDKN--KL 435

Db 340 FKCIIEQVNTMKHEKFAVQDRYTKGQEAITYFITSGGLLLAWLPLTPRSLATVKNLTTEL 399

Qy 436 NNETY-----RRSRQAESTTDPMMEMTNGAGGEYSSENSI 473

Db 400 TPTSPSPSPSPAPSAARGSTPAALVRRRRRDAGNATTP---VPPTAPGKSLGLTLNPP 456

Qy 474 TVAQVAYDNLRIIRNNILEDLSKAWCREQHRALVWNLKSKINPTSVMSMYNRPVSA 533

Db 457 ATVQIOFAYDNLRIIRNNILEDLSKAWCREQHRALVWNLKSKINPTSVMSMYNRPVSA 536

Qy 534 KRIGDVISVNCIVVDQTSVSLHSLRLSASDEKCFSPRPVTPFKMNDSTIYKQGLGVN 593

Db 517 KRLGDVISVOCVPVQATVTLRKSMT--VPGSETMTCYSRPLVSFNFINDTKYEGQLGTD 575

Qy 594 NEILLTTLTLETCOENTYFYFOAKTDMYIKYKNEHLKTVPLSSITLTDITFIALNFTLEN 653

Db 576 NEIFLTKMTKEVCQATSYFYQSGNEIHVYNDVHFKETIELDGIATLQTFISLNTSLIEN 635

Qy 654 VDPKVIETRDEKRLGNVFDIETMFREYNVYQVRVGLSKRLDLD--LSTNRNOFVDAFGS 712

Db 636 IDFASLEYSRDEQASNVFDLEGIHFREYNFQAGNIAGLRKLDLNAVSGNRQFVQDLGE 695

Qy 713 LMDDLGAVGQTVVNAVSGVATLFSSIVTGFINKPFGGMLMIIVVIGVLFALYFLTKK 772

Db 696 LMDSLGVSQSTNLVSTVGGFSSLSVGFISFFKPNPFGMLILVLVAGWILVISLRR 755
Qy 773 TKIYETAPIKMIYPIEDIKLKEREGKSE---IAPISEEELERIVLAMHIHQONSHMETKTR 829
Db 756 TRMSQQPQVMYPIGIDELAQHASGEGPGINPISKTELQAIMLA--LHEQNOBOKRAAQ 813
Qy 830 KDPKDSILTRAQNMRLKR 847
Db 814 RAAGPSVASRALQAADR 831

RESULT 14

US-08-486-099-103
; Sequence 103, Application US/08486099
; Patent No. 6013263
; GENERAL INFORMATION:
; APPLICANT: Bolognesi, Dani P.
; APPLICANT: Matthews, Thomas J.
; APPLICANT: Wild, Carl T.
; APPLICANT: Barney, Shawn O.
; APPLICANT: Lambert, Dennis M.
; APPLICANT: Petteway, Stephen R.
; APPLICANT: Langlois, Alphonse J.
; TITLE OF INVENTION: COMPOSITIONS FOR INHIBITION OF
; TITLE OF INVENTION: MEMBRANE FUSION-ASSOCIATED EVENTS, INCLUDING HEPATITIS
; TITLE OF INVENTION: B VIRUS TRANSMISSION
; NUMBER OF SEQUENCES: 209
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036-2711

COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/486,099
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Coruzzi, Laura A.
; REGISTRATION NUMBER: 30,742
; REFERENCE/DOCKET NUMBER: 7872-031
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-9741/8864
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 103:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 856 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; US-08-486-099-103

Query Match 40.4%; Score 1839; DB 3; Length 856;
Best Local Similarity 46.1%; Pred. No. 1.6e-134; Mismatches 240; Indels 42; Gaps 13;
Matches 368; Conservative 148;

Qy 81 STFPVRVCSAGVGDFRQTDHVCDD-ASDMVHSEGLIYKQNIIPMFVRKRVKVV 139
Db 44 TSFPRVCSLGHDLFRSSDIQCPSTRENHTHGLMWFKDNIIPVSKVRYSYTKIV 103
Qy 140 TTSVTYNGIYSINQHTFYKSIPEWETEKMDTIYQCFNSLRNTGGNLLTYVDRDDIN 199
Db 104 TNLITNGYADSVNTRHEEKFSVDSYETQMDTIYQCYNAVMTKDLTRVYVDRDGVN 163

Qy 200 MTFVLQPDVGVTDPVKRYGSOPELYLEPGFWGWSYRRRTTVCNCLMDMFARSNPPDFV 259
Db 164 ITVNLKPTGGLANGVRVYASQTELYDAPGMLIWTYRTTVCNCLITDMMAKSNPPDFV 223
Qy 260 TATGDTVEMSPWGSDDHKNKMKHFWFVSVINNYKVVDYQNRGTVPGLKTRIFLDREE 319
Db 224 TTTGQTVEMSPFYDGK--NKETPHERADSPHVRVNYKIVDYDNRGTNPQGGERRAFLOKGT 281
Qy 320 YTLSEWKLKN-MSYCPCLTLWKFNGIQTEHSGSYHVFANDITASFTTSKEDMKFNNT 378
Db 282 YTLWK--LENRTAYCPLOHWQTFDSTIATBTGSIHFVDEGTSSPVTNTVIGELPDA 339
Qy 379 YHCLNEEKAEIEKKAIVNSTHSHKYGD-LKYFKTDGGLYLVWQPLQINRLDARN--KL 435
Db 340 FKCIIEQVN-KTHEKYEAVQDRYTKGQEAITYFITSGLLLAWLPLTPRSLATVKNLTTEL 398
Qy 436 NNETY-----RSRRQAESTTDDPMMENTGNGAGEVSSNSI 473
Db 399 TPTSSPPSPSPAPSAARGSTPAVLRRLRRRRDAGNATP---VPPTAPGKSLGTLNPP 455
Qy 474 TVAQVQAYDNLIRINNILEDLSKAWCREOHRALVNMELSKINPTSVMSIYNRPVSA 533
Db 456 ATVQIQAIDSLRQINRMGLDLARACLEOKRQNMVRLTKINPTTVMSSIYKAVAA 515
Qy 534 KRIGDIVSVNCIVVDQTSVSLHSLRLSASDEKCFSPVPTFKFMDNDSTIYKQGLGVN 593
Db 516 KRLGDVLSVQCVPVQATVTLRSMR-VPGSETMYSRPLVSFSPFINDTKTYEGQLGTD 574
Qy 594 NEILLTTLTYLTCQENTYYPQAKTDMYIYKNYKHLKTVPLSSITTLDTFIALNFTLLEN 653
Db 575 NEIFLTKMTEVCQATQYVFGSGNEIHVNDYHFKTIELDGIATLQTFISLNTSLTEN 634
Qy 654 VDEKVIETYDEKRLSNVFDIETMREYNYVAQVSVGLRDKLDD-LSTNRNQFVDFGS 712
Db 635 IDFASLELYKDEQRASNVFDLEGIFREYNFQAQNIAGLRKLDONAVSNRQNFVDGGE 694
Qy 713 LMDDLGAVGQTVNNAVSGVATLFSIVTGFNFINKNPFGGMLMIIVIGVLFAIFYLTKK 772
Db 695 LMDSLGVSQSTNLVSTVGGFSSLSVGFISFFKPNPFGMLILVLVAGWILVISLRR 754
Qy 773 TKIYETAPIKMIYPIEDIKLKEREGKSE---IAPISEEELERIVLAMHIHQONSHMETKTR 829
Db 755 TRMSQQPQVMYPIGIDELAQHASGEGPGINPISKTELQAIMLA--LHEQNOBOKRAAQ 812
Qy 830 KDPKDSILTRAQNMRLKR 847
Db 814 RAAGPSVASRALQAADR 830

RESULT 15

US-08-484-223B-103
; Sequence 103, Application US/08484223B
; Patent No. 6020459
; GENERAL INFORMATION:
; APPLICANT: Bolognesi, Dani P.
; APPLICANT: Matthews, Thomas J.
; APPLICANT: Wild, Carl T.
; APPLICANT: Barney, Shawn O.
; APPLICANT: Lambert, Dennis M.
; APPLICANT: Petteway, Stephen R.
; APPLICANT: Langlois, Alphonse J.
; TITLE OF INVENTION: COMPOSITIONS FOR INHIBITION OF
; TITLE OF INVENTION: MEMBRANE FUSION-ASSOCIATED EVENTS, INCLUDING HIV
; NUMBER OF SEQUENCES: 245
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036-2711
; COMPUTER READABLE FORM:


```
QY 140 TTSTVYNGIYSITNQHTFYKSIPEWETEKMDTIYQCFNSRLRLNTGNNLLTYVDRDDIN 199
Db 104 TNLIIYNGWYADSVNRRHEEKFSVSDYETQMDTIYQCVNAVAKMTDGLTRVYVDRDGVN 163
QY 200 MTVPLOPVDGVTDPVKRYGSOPELYLECGFWGVSRRRTTNCCLMDMFARSNPPDFPV 259
Db 164 ITVNLKPTGGLANGVRRYASOTELYDAPGWLITVYTRTTVNCCLITDMMKSNPSDFPV 223
QY 260 TATGDTVEMSPFSGEDDHNKMKHKEKPFVSVINNYKVVDYQNRGTVPGLKTRIFLDREE 319
Db 224 TTGTQTVEMSPFYDGK--NKETPHERADSFHVRTNYKIVDYNRGTNPQGERRAFLDKGT 281
QY 320 YTLSEKHLKN--MSYCPBLTWKAFYNGIOTHSYSHFVANDITASFTTSKEDMKFNTT 378
Db 282 YTLSEK--LENRTAYCPLQHWQTFDSTIATETGKSIHFVDEGTSFVNTTVGIELPDA 339
QY 379 YHCLNEEKAEIEKKYAKVNSHSHSKYGD--LKFKTDGGLYLVWQPLIQNRLLDKNN--KL 435
Db 340 FKCIIEQVN--KTHEKYEAVQDRYTKGQEAITYFTISGGLLAWLPLTPRSLATVKNLTTEL 398
QY 436 NNETYSS-----RRSRQAESTTDPMMEMTNGAGGEYSSENSI 473
Db 399 TPTSPSPSPSPAPSAARGSTPAAVLRRRRRRDAGNATTP---VPPTAPGKSLGTNNP 455
QY 474 TVAQVQYAYNLRIRINNILEDLKAWCREQHRALVNNELSKINPTSVMSIYNRPVSA 533
Db 456 ATVOIQFAYDSLRRQINRMGLDLARAWCLEQKQNMVLRRLTKINPTTVMSSIYKAVAA 515
QY 534 KRIGDVISVNCIVVDQTSVLSHLKSLRLLSASDEKCFRPPVTFKFMNDSTIYKQGLGVN 593
Db 516 KRLGDVISVQCVPVNOATVTLRKSMR--VPGSETMCSYRPLVSFINDTKTYEGQLGTD 574
QY 594 NEILLTTLTLETCEQNTYVFOAKDMYIKYVHKTVPVLSITLDTFIALNFTLLEN 653
Db 575 NEIFLTKMTEVCQATSYQFOSGNEIHVNDYHFKTIELDGIATLQTFISLNTSLIEN 634
QY 654 VDFKVIELYTRDEKRLSNVFDIETMFREYNYYAQRVSGRLKDLDD--LSTNRNQFVDAFGS 712
Db 635 IDFASLELYSRDEQASNVDFLEGIFREYNFQAQNIAGLRKDLDNVNSGRNQFVDGLGE 694
QY 713 LMDDLGAVQTVNAVSGVATLFFSSIVTGFNFINKPFGGMLIIVIGVLFAIYFLTKK 772
Db 695 LMDLSGVSQGITNLVSTVGLFSSLSVSGFISFFKNPFGGMLILVLVAGVILVISLPR 754
QY 773 TKIYETAPIKMYPIBDIKKREGKSE---TAPISEEELERIVLAWHTHOQNSHMETKTR 829
Db 755 TRQMSQQVQMLYPCIDELAQHASGEGPGINPISKTELQAIMLA--LHEQNEQOKRAAQ 812
QY 830 KDPKDSILTRAQNMLRKR 847
Db 813 RAAGPSVASRALQAADR 830
```

RESULT 17

```
US-08-475-668A-103
; Sequence 103, Application US/08475668A
; Patent No. 6060065
;
; GENERAL INFORMATION:
; APPLICANT: Barney, Shawn O.
; APPLICANT: Lambert, Dennis M.
; APPLICANT: Pattaway, Stephen R.
; TITLE OF INVENTION: COMPOSITIONS FOR INHIBITION OF MEMBRANE
; TITLE OF INVENTION: FUSION-ASSOCIATED EVENTS, INCLUDING INFLUENZA VIRUS
; TITLE OF INVENTION: TRANSMISSION
; NUMBER OF SEQUENCES: 211
; CORRESPONDENCE ADDRESS:
; ADDRESS: Pennie & Edmonds LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
```

```
;
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/475.668A
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Cortuzzi, Laura A.
; REGISTRATION NUMBER: 30,742
; REFERENCE/DOCKET NUMBER: 7872-026
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-9741/8864
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 103:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 856 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; US-08-475-668A-103
```

```
Query Match 40.4%; Score 1839; DB 3; Length 856;
Best Local Similarity 46.1%; Pred. No. 1.6e-134;
Matches 368; Conservative 148; Mismatches 240; Indels 42; Gaps 13;
QY 81 STFPYRVCSAGVGDFRFOTDHYCPD--ASDMVHSEGLLIYKQNIIPFMRVRKRYKV 139
Db 44 TSFPFRVCELSSHGDLFRFSSDIQCPSTRENHTEGLLMVFKDNIIPYSFKVRSYTKIV 103
QY 140 TTSTVYNGIYSDSTNQHTFYKSIPEWETEKMDTIYQCFNSRLRLNTGNNLLTYVDRDDIN 199
Db 104 TNLIIYNGWYADSVNRRHEEKFSVSDYETQMDTIYQCVNAVAKMTDGLTRVYVDRDGVN 163
QY 200 MTVPLOPVDGVTDPVKRYGSOPELYLECGFWGVSRRRTTNCCLMDMFARSNPPDFPV 259
Db 164 ITVNLKPTGGLANGVRRYASOTELYDAPGWLITVYTRTTVNCCLITDMMKSNPSDFPV 223
QY 260 TATGDTVEMSPFSGEDDHNKMKHKEKPFVSVINNYKVVDYQNRGTVPGLKTRIFLDREE 319
Db 224 TTGTQTVEMSPFYDGK--NKETPHERADSFHVRTNYKIVDYNRGTNPQGERRAFLDKGT 281
QY 320 YTLSEKHLKN--MSYCPBLTWKAFYNGIOTHSYSHFVANDITASFTTSKEDMKFNTT 378
Db 282 YTLSEK--LENRTAYCPLQHWQTFDSTIATETGKSIHFVDEGTSFVNTTVGIELPDA 339
QY 379 YHCLNEEKAEIEKKYAKVNSHSHSKYGD--LKFKTDGGLYLVWQPLIQNRLLDKNN--KL 435
Db 340 FKCIIEQVN--KTHEKYEAVQDRYTKGQEAITYFTISGGLLAWLPLTPRSLATVKNLTTEL 398
QY 436 NNETYSS-----RRSRQAESTTDPMMEMTNGAGGEYSSENSI 473
Db 399 TPTSPSPSPSPAPSAARGSTPAAVLRRRRRRDAGNATTP---VPPTAPGKSLGTNNP 455
QY 474 TVAQVQYAYNLRIRINNILEDLKAWCREQHRALVNNELSKINPTSVMSIYNRPVSA 533
Db 456 ATVOIQFAYDSLRRQINRMGLDLARAWCLEQKQNMVLRRLTKINPTTVMSSIYKAVAA 515
QY 534 KRIGDVISVNCIVVDQTSVLSHLKSLRLLSASDEKCFRPPVTFKFMNDSTIYKQGLGVN 593
Db 516 KRLGDVISVQCVPVNOATVTLRKSMR--VPGSETMCSYRPLVSFINDTKTYEGQLGTD 574
QY 594 NEILLTTLTLETCEQNTYVFOAKDMYIKYVHKTVPVLSITLDTFIALNFTLLEN 653
Db 575 NEIFLTKMTEVCQATSYQFOSGNEIHVNDYHFKTIELDGIATLQTFISLNTSLIEN 634
QY 654 VDFKVIELYTRDEKRLSNVFDIETMFREYNYYAQRVSGRLKDLDD--LSTNRNQFVDAFGS 712
Db 635 IDFASLELYSRDEQASNVDFLEGIFREYNFQAQNIAGLRKDLDNVNSGRNQFVDGLGE 694
```

Qy 713 LMDLGVAGVTVNAVSGVATLFPSSIVTGFNFIKPNFGMLMIIVVIGVLFALYFLTKK 772
Db 695 LMDSLGSGVGSITNLVSTVGGFLSSLVSGRISFFKPNFGMLLVLVAGVVLVISTRR 754
Qy 773 TKIYETAPIKMIYPIEIDKLKREKSE---IAPISEEELERIVLAMHIHOONSHMETKTR 829
Db 755 TRQMSQOPVQMLYPGIDELAQOAHASGEGPGINPISTKTELQAIMLA--LHEQNOEQKRAAQ 812
Qy 830 KDPKDSILTRAQNMRLKR 847
Db 813 RAAGPSVASRALQAARDR 830

RESULT 18
US-08-485-551A-103
; Sequence 103, Application US/08485551A
; Patent No. 6068973
; GENERAL INFORMATION:
; APPLICANT: Bolognesi, Dani P.
; APPLICANT: Matthews, Thomas J.
; APPLICANT: Wild, Carl T.
; APPLICANT: Barney, Shawn O.
; APPLICANT: Lambert, Dennis M.
; APPLICANT: Petteway, Stephen R.
; APPLICANT: Langlois, Alphonse J.
; TITLE OF INVENTION: METHODS FOR INHIBITION OF MEMBRANE
; TITLE OF INVENTION: FUSION-ASSOCIATED EVENTS, INCLUDING INFLUENZA VIRUS
; NUMBER OF SEQUENCES: 211
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/485,551A
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Coruzzi, Laura A.
; REGISTRATION NUMBER: 30,742
; REFERENCE/DOCKET NUMBER: 7872-023
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-9741/8864
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 103:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 856 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; US-08-485-551A-103

Query Match 40.48; Score 1839; DB 3; Length 856;
Best Local Similarity 46.18; Pred. No. 1.66-134;
Matches 368; Conservative 148; Mismatches 240; Indels 42; Gaps 13;

Qy 81 STFPYRVCSAGVDVRFOTDHFVCPD-ASDMVHSEGILLIYKQNIIPFMFRVRYKRVV 139
Db 44 TSFPFRVCELSSHGDLPFRFSDLOCPSGFGTRENHTEGLLMVFKDNIIPYFKVRSYKIV 103
Qy 140 TTSTVYNGIVSDSITNOHTYKSIPEWETEKMDTIYOCFNSRLNTGCLLLTVYDRDIN 199
Db 104 TNLIIYNGWYADSVNTRHEEKFSVDSYETDQMDTIYOCYNVAVKMTDGLTRVYVDRDGN 163

Qy 200 MTFVLOPVDGVTEPDKVYGSQPELYLEPGWFGSYRRRTTVNCELMDMPARSPPPDFV 259
Db 164 ITVNLKPTGGLANGVRYASQTELYDAPGLWIWYTRITVNCCLITDMMAKSPDPDFV 223
Qy 260 TATGDTVEMSPWMSGDDHENKMKHPWFWFVSVNNYKVDYQNRGTVPGLGKTRIFLDREE 319
Db 224 TTTGQIVEMSPFDYDK--NKETPHERADSFHVRTNYKIVDYDNRGTINPOGERAFDLKGT 281
Qy 320 YTLSEKHLKN-MSYCPLTLWKAFYNGIOTEHSGSYHFVANDITASTFTSKEDMKENFTT 378
Db 282 YTLNWK--LENRTAYCPLQHWOTFDSTIATETGKSIHFVTDEGTSFVNTTVGIELPDA 339
Qy 379 YHCLNEEIKAEIEKYAKYNSTHSHKYGD-LKYFKTDGGLYLVWQPLIONRLLDKKN--KL 435
Db 340 FCIEEQVN-KTHEKYEAQVDRYKQEAITYFITSGLLLAWLPLTPRSLATVKNLTEL 398
Qy 436 NNETYS-----RRSRQAESTTDPMMEMTNGAGGEYSSENSI 473
Db 399 TPTSPSPSPSPAPSAARGSTPAAVLRRRRDAGNATP---VPPTAPGKSLGTLNPP 455
Qy 474 TVAQOYAYDNLRIRINNILEDLSKAWCREQHRALVWNELSKINPTSVMSMIYNRPVSA 533
Db 456 ATVQIQFAYDSLRRQINRMGLDLARAWCLEQKRNQNVMLRELTKINPTVMSSTYGVKAAV 515
Qy 534 KRIGDVISVNCIIVDQTSVSLHKSRLLSASDEKCFSPVPVTFKEMNDSTIYKGGOLGVN 593
Db 516 KRIGDVISVQCVPVNVQATVTLKSMR-VPGSETMCSRPLVSFSFINDTKYEGQLGTD 574
Qy 594 NEILLTTLTTCQENTYFYFOAKTDMYIKNYEHLKTVPLSSITTLDTFIALNFTLEN 653
Db 575 NEIFLTKMTEVQATSOYVYFQSGNEIHVYNDYHHEKTIELDGIATLQTFISLNTSLIEN 634
Qy 654 VDFKVIETRYDEKRLSNVFDIETMPREYNVYVAQRVSGLRKDLDD--LSNRRNQFVDAFGS 712
Db 635 IDFASLELYSRDEQRASNVFDLEGIPEYFNFOAQNIAGLRKOLDNAVSGNRQNVFDGLGE 694
Qy 713 LMDLGVAGVTVNAVSGVATLFPSSIVTGFNFIKPNFGMLMIIVVIGVLFALYFLTKK 772
Db 695 LMDSLGSGVGSITNLVSTVGGFLSSLVSGRISFFKPNFGMLLVLVAGVVLVISTRR 754
Qy 773 TKIYETAPIKMIYPIEIDKLKREKSE---IAPISEEELERIVLAMHIHOONSHMETKTR 829
Db 755 TRQMSQOPVQMLYPGIDELAQOAHASGEGPGINPISTKTELQAIMLA--LHEQNOEQKRAAQ 812
Qy 830 KDPKDSILTRAQNMRLKR 847
Db 813 RAAGPSVASRALQAARDR 830

RESULT 19
US-08-471-913A-103
; Sequence 103, Application US/08471913A
; Patent No. 6093794
; GENERAL INFORMATION:
; APPLICANT: Bolognesi, Dani P.
; APPLICANT: Matthews, Thomas J.
; APPLICANT: Wild, Carl T.
; APPLICANT: Barney, Shawn O.
; APPLICANT: Lambert, Dennis M.
; APPLICANT: Petteway, Stephen R.
; APPLICANT: Langlois, Alphonse J.
; TITLE OF INVENTION: COMPOSITIONS FOR INHIBITION OF MEMBRANE
; TITLE OF INVENTION: FUSION-ASSOCIATED EVENTS, INCLUDING EPSTEIN-BARR VIRUS
; NUMBER OF SEQUENCES: 214
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036-2711

[illegible]

Db 635 IDFALELYSRDQRASNVFDLEGIREFYNFOAQNAGRLKOLDNAVSNRQFVDGLGE 694
 QY 713 LMDLGAAGQTVNNAVSGVATLFSISIVTGFINKPFGMLMIIVIGLFAIYFLTKK 772
 Db 695 LMDSLGVSQISYINLVSTVGGFESSLVSGFISFKPFGMLLVLVAGVILVISLTKR 754
 QY 773 TKIYETAPIKMIYPEIDKLKEREKSE---IAPISEEELERIVLAMHIHQSNHMETKTR 829
 Db 755 TROMSQOPVQMLYPGIDELAQHASGEGPGINPISKTELQAIMLA--LHEQNEQOKRAAQ 812
 QY 830 KDPKDSILTRAQNMRLKR 847
 Db 813 RAAGPSVASRALQAARDR 830

RESULT 22
 US-08-470-896-103
 ; Sequence 103, Application US/08470896
 ; Patent No. 6479055
 ; GENERAL INFORMATION:
 ; APPLICANT: Bolognesi, Dani P.
 ; APPLICANT: Matthews, Thomas J.
 ; APPLICANT: Wild, Carl T.
 ; APPLICANT: Barney, Shawn O.
 ; APPLICANT: Lambert, Dennis M.
 ; APPLICANT: Petteway, Stephen R.
 ; APPLICANT: Langlois, Alphonse J.
 ; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR INHIBITION
 ; TITLE OF INVENTION: OF MEMBRANE FUSION-ASSOCIATED EVENTS, INCLUDING HIV
 ; NUMBER OF SEQUENCES: 273
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Pennie & Edmonds
 ; STREET: 1155 Avenue of the Americas
 ; CITY: New York
 ; STATE: New York
 ; COUNTRY: USA
 ; ZIP: 10036-2711
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent In Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/470,896
 ; FILING DATE: 06-JUN-1995
 ; CLASSIFICATION: 435
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Coruzzi, Laura A.
 ; REGISTRATION NUMBER: 30,742
 ; REFERENCE/DOCKET NUMBER: 7872-020
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (212) 790-9090
 ; TELEFAX: (212) 869-9741/8864
 ; TELEX: 66141 PENNIE
 ; INFORMATION FOR SEQ ID NO: 103:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 856 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS:
 ; TOPOLOGY: unknown
 ; MOLECULE TYPE: protein
 ; US-08-470-896-103

Query Match 40.4%; Score 1839; DB 4; Length 856;
 Best Local Similarity 46.1%; Pred. No. 1.6e-134;
 Matches 368; Conservative 148; Mismatches 240; Indels 42; Gaps 13;

QY 81 STFPYVCSASGVDFRQTHVCPD-ASDMVHSGILLIYKQNIIPMPFRVRYKVV 139
 Db 44 TSFPFRVCSLSHGDLFRFSSDIQCFSTRENHTGLMVPKDNIIIPYSFKVRSYTKIV 103
 QY 140 TTSTVNGIYSDISITNQHTFYKSIPEWTERKMDTIYQCFNSURLNTGGNLLTYVDRDDIN 199

Db 104 TNILYNGWYADSVTNRHEEKFSVDSYETDQMDTIYQCNVAKMTKDCGLTRVYVDRDGVN 163
 QY 200 MTVFLOPVDGVTDPVKRYGSGPELYLEPGWFGWYSRRRTTNCBLMDMFASNPFFDFV 259
 Db 164 ITVNLKPTGGGLANGVRRYASQTLEYDAPGMLIWTYRTRTTTNNCLITDMMAKSNSPFFDFV 223
 QY 260 TATGDTVEMSPFWSGEDDHENKHEKPMFVSVINNYKVVDYQNRGTVPGLGTRIFLDREE 319
 Db 224 TTTGQTVEMSPFYDGK--NKETFFERADS FVRYNYKIYVDYDNRGTNPQGERAFLDKGT 281
 QY 320 YTLSEKHLKN--MSYCPLTLKAFYNGIQTEHSGSYHFVANDITASFTTSKEDKKEFNTT 378
 Db 282 YTLGK--LENRTAYCPLQHWQTDSTIATETGKSIHEVTDEGTSSFVNTTGVIELPDA 339
 QY 379 YHCLNEEIKAEIKYAKVNSTHSHKYGD-LKYFTDGLGLYLWQPLIQNRLLDAKN--KL 435
 Db 340 FKCIIEQVN-KTHEKYEAVQDRYTKGQEAITYFTTSGGLLLAWLPLTPRSLATVKNLTTEL 398
 QY 436 NNETYS-----RRSRQAESTTDPMMEMTGNAGGEYSSENSI 473
 Db 399 TPTSSPPSPSPAPPASARGSTPAAVLRRRRRRDAGNATTP---VPPTAPKSLGTLNPP 455
 QY 474 TVAQVQYAYDNLRIINNILEDLSKAMCREQRAALVWNELSKINPTSVMSMIYNRPVSA 533
 Db 456 ATVOIQPAYDSLRRQINRMGLDLARAWCLEQKQNMVLRLELTKINPTTVMSSIVGKAVAA 515
 QY 534 KRIGDIVISVNCIVVDQTSVLHKSLLLSASDEKCSRPPVTPKFMNDSTIYKGOLGVN 593
 Db 516 KRLGDVIVSVQCVPVNAQTATVLRKSMR--VPGSETMCSYRPLVSVSFINDTITYEQQLGTD 574
 QY 594 NEILLTTTYLETCEONTYEFYFOAKTDMYIYKNYEHKLTVPPLSSITTLDTFALNFTLLEN 653
 Db 575 NEIFLTWKTEVCQATSOYIFQSGNEIHVYNDYHHFTIELDGIATLQTFISLNTSLTEN 634
 QY 654 VDFKVIELYTRDEKRLSNVFDIETMFREYNYYAQRVSLRGLRLDLD--LSTNRNQFYDARFS 712
 Db 635 IDFALELYSRDEQRASNVFDLEGIREFYNFOAQNAGRLKOLDNAVSNRQFVDGLGE 694
 QY 713 LMDLGAAGQTVNNAVSGVATLFSISIVTGFINKPFGMLMIIVIGLFAIYFLTKK 772
 Db 695 LMDSLGVSQISYINLVSTVGGFESSLVSGFISFKPFGMLLVLVAGVILVISLTKR 754
 QY 773 TKIYETAPIKMIYPEIDKLKEREKSE---IAPISEEELERIVLAMHIHQSNHMETKTR 829
 Db 755 TROMSQOPVQMLYPGIDELAQHASGEGPGINPISKTELQAIMLA--LHEQNEQOKRAAQ 812
 QY 830 KDPKDSILTRAQNMRLKR 847
 Db 813 RAAGPSVASRALQAARDR 830

RESULT 23
 US-08-485-546A-103
 ; Sequence 103, Application US/08485546A
 ; Patent No. 6518013
 ; GENERAL INFORMATION:
 ; APPLICANT: Bolognesi, Dani P.
 ; APPLICANT: Matthews, Thomas J.
 ; APPLICANT: Wild, Carl T.
 ; APPLICANT: Barney, Shawn O.
 ; APPLICANT: Lambert, Dennis M.
 ; APPLICANT: Petteway, Stephen R.
 ; APPLICANT: Langlois, Alphonse J.
 ; TITLE OF INVENTION: METHODS FOR INHIBITION OF MEMBRANE
 ; TITLE OF INVENTION: FUSION-ASSOCIATED EVENTS, INCLUDING EPSTEIN-BARR VIRUS
 ; NUMBER OF SEQUENCES: 214
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Pennie & Edmonds LLP
 ; STREET: 1155 Avenue of the Americas
 ; CITY: New York
 ; STATE: New York

Db 136 TGKFMATQAEHWEVDFDSIIQCNVSNATMVNNRVQVYVDRDGVNKTINIRPVDGLTGN 195
Qy 214 VKRYGSOPELYLEPGWFGWSYRRRTTVNCELMDMFARSNPPDFVVTATGDTVEMSPFWS 273
Db 196 IQRYFSQPTLYSEPGMPGFYRVRTVNCIEIDMVARSMDPNYIATAGDLSLESPFOT 255
Qy 274 GEDDHENKMKHEK-PWFVSVINNYKVVDYQNRGTVPGLKTRIFLDREBEYTLSEKHLKMS 332
Db 256 FNTSQCTAPKADMRVREVKVYFVDYNNRGTAPAGOSRTFLETSPATYSWKATATQTA 315
Qy 333 YCPILTAKFAYNGIQTEHSGSHYFVANDITAGFTTSKEDMKFNTTYHCLNEIEKAEIEK 392
Db 316 TCDLVHMKTFPRAIQTAHSHYFVANEVATFNTPLTEVENFTSYSCVSDQINKTISE 375
Qy 393 KYAKVNSTHSGYGLDKYFKTGGGLVLMQPLQNRLLDKKNLNE-----TYSRRSRQ 447
Db 376 YIQKLNNSYVASKGTQYFKTDGNLYLIWQPLEHEIEDIDEDSDPEPTAPPKSTRKRE 435
Qy 448 AESTTDPMMMTGNGAGGEYSSENSITVAQVQYAYDNLRIRINNILEDLKAWCREQHRA 507
Db 436 AADNGNSTSEVS-----KGSENPLITAQIQFAYDKLTTSVNNVLEELSRACREQVRD 488
Qy 508 ALVWELSKINTSVMSMTYNRPVSAKRIGDVISVNCIIVDQTSVSLHKSURLLSASDE 567
Db 489 TLMWYELSKVNPSTVMSAIYKGPVAAARYGDAISVTDICIYVDQSSVNIHQSLR-LQHDKT 547
Qy 568 KCFSRPPVTFKFMNDSTIYKGGOLGVNNEILLTTTYLETQCENTEYFQAKTDMYIKNYE 627
Db 548 TCYSRPRVTFKFINSTDPGLTQGLGPRKEIILSNINIECTKDESEHYFVGEVYIYKNYI 607
Qy 628 HLKTVPLSSITLDTFIALNFTLLENVDFKVELYTRDEKRL-SNVFDETFRFVNYA 686
Db 608 FEEKLNLSIATLDTFIALNISFIENIDFKVELSSYTERKCLASSVDFIESMFRFNYIT 667
Qy 687 QRVSLRKOLL-D-LSTNRNQFVDAFGSLMDLGAQVQTVVNAVSGVATLFFSSIVTGFINF 745
Db 668 YSLAGIKKOLDNTIDYNRDLVQDLSDMDLGDIGRSVNVVSSVTFSSIVTGFIKF 727
Qy 746 IKNPFGGMLMIIVVIGLVFAIYFLTKTIYETAPIKMIYPEIDKLRREG-----KSEIA 801
Db 728 FNPJGGFILLIIGIIFLVVLAARRNSQFHDPIKMLYPSVENYAAQAPPPYSASPP 787
Qy 802 PISEBELERIVLAME-IHQONSHMETK-TRKDPKDSILTRAQNMLR-KRSGYSNL 853
Db 788 AIDKEIKRILLGMQVHQEKEAQKLTNSGP--TLWQKATGFLNRKKGISQL 840

RESULT 25

US-08-720-229-17
Sequence 17, Application US/08720229
Patent No. 6022542
GENERAL INFORMATION:
APPLICANT: Rose, Timothy M.
APPLICANT: Bosch, Marnix L.
APPLICANT: Strand, Kurt
TITLE OF INVENTION: GLYCOPROTEIN B OF THE RBHV/KSHV
SUBFAMILY OF HERPES VIRUSES
NUMBER OF SEQUENCES: 100
CORRESPONDENCE ADDRESS:
ADDRESSEE: Morrison & Foerster
STREET: 755 Page Mill Road
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304-1018
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/720,229

FILING DATE: 26-SEP-1996
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Schieff, J. Michael
REGISTRATION NUMBER: 40,253
REFERENCE/DOCKET NUMBER: 29938-20002.00
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 813-5600
TELEFAX: (415) 494-0792
TELEX: 706141
INFORMATION FOR SEQ ID NO: 17:
SEQUENCE CHARACTERISTICS:
LENGTH: 849 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-720-229-17
Query Match 39.4%; Score 1792; DB 3; Length 849;
Best Local Similarity 42.4%; Pred. No. 7, 2e-131;
Matches 354; Conservative 158; Mismatches 271; Indels 52; Gaps 13;
Qy 34 SPPNTATWSTSPLTGCHYTHDSHGGRNENRDSSEQNKNIYGPSSTFPYRVCSAGV 93
Db 43 TPAQDAPTETPPPLS-----TWTNRGFY-----PRVCGVAAT 75
Qy 94 GDVFRFQDHCVPDASDMHSEGLLIYKQMIIPFMRVRKYRKVVTTSTVYNGIYSDSI 153
Db 76 GETFRFDLDKTCPSQDQKKHVEGILLVYKINIVPIFKIRRYRKIIITQITWRGLTSSV 135
Qy 154 TNOHTFYKSIPEWETKMDTLTYQCFNSLRLATGNNLLTYVDRDDINMTVFLQPVGVTPD 213
Db 136 TCKEMATQAEHWEVDFDSIIQCNVSNATMVNNRVQVYVDRDGVNKTINIRPVDGLTGN 195
Qy 214 VKRYGSOPELYLEPGWFGWSYRRRTTVNCELMDMFARSNPPDFVVTATGDTVEMSPFWS 273
Db 196 IQRYFSQPTLYSEPGMPGFYRVRTVNCIEIDMVARSMDPNYIATAGDLSLESPFOT 255
Qy 274 GEDDHENKMKHEK-PWFVSVINNYKVVDYQNRGTVPGLKTRIFLDREBEYTLSEKHLKMS 332
Db 256 FNTSQCTAPKADMRVREVKVYFVDYNNRGTAPAGOSRTFLETSPATYSWKATATQTA 315
Qy 333 YCPILTAKFAYNGIQTEHSGSHYFVANDITAGFTTSKEDMKFNTTYHCLNEIEKAEIEK 392
Db 316 TCDLVHMKTFPRAIQTAHSHYFVANEVATFNTPLTEVENFTSYSCVSDQINKTISE 375
Qy 393 KYAKVNSTHSGYGLDKYFKTGGGLVLMQPLQNRLLDKKNLNE-----TYSRRSRQ 447
Db 376 YIQKLNNSYVASKGTQYFKTDGNLYLIWQPLEHEIEDIDEDSDPEPTAPPKSTRKRE 435
Qy 448 AESTTDPMMMTGNGAGGEYSSENSITVAQVQYAYDNLRIRINNILEDLKAWCREQHRA 507
Db 436 AADNGNSTSEVS-----KGSENPLITAQIQFAYDKLTTSVNNVLEELSRACREQVRD 488
Qy 508 ALVWELSKINTSVMSMTYNRPVSAKRIGDVISVNCIIVDQTSVSLHKSURLLSASDE 567
Db 489 TLMWYELSKVNPSTVMSAIYKGPVAAARYGDAISVTDICIYVDQSSVNIHQSLR-LQHDKT 547
Qy 568 KCFSRPPVTFKFMNDSTIYKGGOLGVNNEILLTTTYLETQCENTEYFQAKTDMYIKNYE 627
Db 548 TCYSRPRVTFKFINSTDPGLTQGLGPRKEIILSNINIECTKDESEHYFVGEVYIYKNYI 607
Qy 628 HLKTVPLSSITLDTFIALNFTLLENVDFKVELYTRDEKRL-SNVFDETFRFVNYA 686
Db 608 FEEKLNLSIATLDTFIALNISFIENIDFKVELYSSYTERKCLASSVDFIESMFRFNYIT 667
Qy 687 QRVSLRKOLL-D-LSTNRNQFVDAFGSLMDLGAQVQTVVNAVSGVATLFFSSIVTGFINF 745
Db 668 YSLAGIKKOLDNTIDYNRDLVQDLSDMDLGDIGRSVNVVSSVTFSSIVTGFIKF 727
Qy 746 IKNPFGGMLMIIVVIGLVFAIYFLTKTIYETAPIKMIYPEIDKLRREG-----KSEIA 801

Db 728 FTNPLGGIFILLIIGIIFLVVLLNRRNSQPHDAPIKMLYPSVNYAARQAPPYSASPP 787
Qy 802 PISEEELERIVLHM-IHQONSHMETK-TRKDPKDSILTRAQNMRL-KRSGYSNL 853
Db 788 AIDKEEIKRILLGHQVHQBEEAKQKLTNSGP--TLWQKATGFLRNRKGYSQL 840
RESULT 26
US-08-804-439A-15
; Sequence 15, Application US/08804439A
; Patent No. 6015565
; GENERAL INFORMATION:
; APPLICANT: Rose, Timothy M.
; APPLICANT: Bosch, Marnix L.
; APPLICANT: Strand, Kurt
; TITLE OF INVENTION: GLYCOPROTEIN B OF THE RFHV/KSHV
; TITLE OF INVENTION: SUBFAMILY OF HERPES VIRUSES
; NUMBER OF SEQUENCES: 113
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 4225 Executive Square, Ste 1400
; CITY: La Jolla
; STATE: CA
; COUNTRY: USA
; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/804.439A
; FILING DATE: February 21, 1997
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Haile, Lisa A.
; REGISTRATION NUMBER: 38,347
; REFERENCE/DOCKET NUMBER: 09176/004001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 678-5070
; TELEFAX: (619) 678-5099
; TELEX:
; INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 874 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-804-439A-15
Query Match 39.3%, Score 1787; DB 3; Length 874;
Best Local Similarity 43.5%; Pred. No. 1.8e-130;
Matches 367; Conservative 159; Mismatches 248; Indels 70; Gaps 20;
Qy 34 SPNTATWSTESPLTGHGTHDSHGGRNENRDSQKNIIYGSSTPYRVCSA-SG 92
Db 58 SPQNT--STSKPSTDNQGSTPT-----IPTVDDTASKNFY-----KYRVCSSSS 102
Qy 93 VGDVFRFQTHVCPDASDMVHSEGLIILYKONIIIPFMFRVKRYKVVTTSTVYNGIYSDS 152
Db 103 SGELFRFDLQTCPTDKDKKHVEGILLVLLKKNIVPIYFKVKRYKIAKTSVTVYRGSQAA 162
Qy 153 ITNQHTYKIEPHEMETKMDTIYCFNSLRNLNTGNNLLTVYDRDDINMTFLQPDGVTP 212
Db 163 VTNRDDISRAIPYNEISMIDRTYHCFSAMATVINGILNTYIDRDSENKSVLPQFVAGLTE 222
Qy 213 DVKRYGSOPELXLPFGWGSYRRTTVNCELMDMFARSNPPDPFFVTATGDTVEMSPFW 272
Db 223 NINRYFSQPLIYABPGHFGPIYRVTVVCEVDMYARSVEPYTHFITALGDTIEISPPC 282
Qy 273 SGEDDHENKM-----HEKPFVSVINNYKVVDYQNRGTVPGLKTRIFLDREYTL 322

Db 283 -----HNNSOCTGNSTSRDATKVV---IBENHQTVDYERRGH-PTKDKRIFLKDEEYTI 333
Qy 323 SWEKHLKNMSYCPILTLWKAIFYNGIOTSEHSYHFVANDITASFTTSKEDMKEP--NT-TY 379
Db 334 SWAEDRERAICDFVIMKTFPRAIQTIHNSPFHFVANEVASFUTSNOEBETELRGNTAIL 393
Qy 380 HCLNEEIKABIEKKYAKVNSTHSHSKYGLKFKYKTDGGLYLVQPLIQNRLDANK---KLN 436
Db 394 NCWNSTINLEETVTKFKNSHIRDGEVKYKTKNGGLFLIWOAKMPLNLSHNTYTIERN 453
Qy 437 NETYSRRSRQABSTTDPMMEMTGNGAGGYSSENSITVAQVOYAYDNLRIINILEDL 496
Db 454 NKTGNKSRQKRSVDT-----KTFQKAG-----LSTAQVOYAYDHLATSMNHILEEL 500
Qy 497 SKAWCREQHPAALVWNLKSKINPTSVMSMIYNRPVSAKRIGDVISVNCIIVDQTSYSLH 556
Db 501 TKTWCREQKDNLMWYELSKINPVSVMAAIYGRKPAVKMGDAFMVSECINVDQASVNIH 560
Qy 557 KSLRLLSASDEK-CFSRPPVTFKPMNDSTIYKQGLGVNNEILLTTTLYLETQOENTEYFQ 615
Db 561 KSMR---TDDPKVCYSRPLVTFKPVNSTATFRQLGTRNEILLTNTHVEICRPTADHYFF 617
Qy 616 AKTDMYIKNYEHLKTVPLSSITTLDTFIALNFTLLENVDFKVIETYRDEKLSNVFDI 675
Db 618 VKMNTYFKDYKFKVKTMDTNNISTLDTFLTLNLTDFIDNIDFKTVELYSETERKVASALDL 677
Qy 676 ETMFREYNYAQRVSGLRKDL---LDLSTNRNPFVDAFGLMDLGLAGVGTGVNAVSGVA 732
Db 678 ETMFREYNYTQKIASLRDLNLTIDL--NRDLVKDLSEMMADLGDIGKVVVNTFSGIV 735
Qy 733 TLFSSIVTGFINFINKPFGGMLMIIVVIGVLPFIYFLTKTKIYETAPIKMIYPIEDKLK 792
Db 736 TVFGSIVGGFVSFTNPICGVITILLIIVVVFVIVSRRTNNNEAPIKMIYPIEDKAS 795
Qy 793 EREGKSEIAPISEEELERIVLAVHHHQONSHMETKTRKDPK---SILTRAQNMRLKRSG 849
Db 796 EQE---NIQPLPGEIKRILLGMHQLQOSEHGKSEEBASHKPGLFQLLGLDGLQLLRRR-G 851
Qy 850 YSNL 853
Db 852 YTRL 855
RESULT 27
US-08-720-229-15
; Sequence 15, Application US/08720229
; Patent No. 6022542
; GENERAL INFORMATION:
; APPLICANT: Rose, Timothy M.
; APPLICANT: Bosch, Marnix L.
; APPLICANT: Strand, Kurt
; TITLE OF INVENTION: GLYCOPROTEIN B OF THE RFHV/KSHV
; TITLE OF INVENTION: SUBFAMILY OF HERPES VIRUSES
; NUMBER OF SEQUENCES: 100
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morrison & Foerster
; STREET: 755 Page Mill Road
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304-1018
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/720,229
; FILING DATE: 26-SEP-1996
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Schiff, J. Michael
; REGISTRATION NUMBER: 40,253

	Query Match	24.8%; Score 1128; DB 1; Length 907;
	Best Local Similarity	30.9%; Pred. No. 3.9e-79;
	Matches	286; Conservative 184; Mismatches 353; Indels 102; Gaps 25;
Qy	14 LWTLYQVALYSLSIATG--VTSPPNTATWSTESPLTGHGTIDSS--HGERGNENRDS	69
	: : :	
Dd	5 IWCL--VVCNLCIVCLGAUVSSSTRGTSAHS---HHSHTTSAHSRSGSVSRVT	58
	: : :	
Qy	70 EQ-----NKNIYG-----SPSFPFYRCVSASGVDFRQTDHVC----PDA	108
	: : :	
Dd	59 SSQTVSHGVNETIYTTLTKYGDVVGVNTTKPYPRCVCSMAOGLDLIRFERINVTSMKPIN	118
	: : :	
Oy	109 SDMVHSEGILLIKONIIPEMFRVKRYKVVTYSTVNGIYSISITKHOFYFKSIERPET	168
	: : :	

Db 119 EDL--DEGIMVVKRNIVAHFTKVRVQKVLTRRSVAYIHTTYLLGSNTSEYVAPPWMEI 176
Qy 169 EKMDTIYOCNSLRNLTTGNNLLTVVDRDDI--NMTVFLQPVGDVTPDVKRYGSOPELYLEP 227
Db 177 HHINSHQCYSSYSSRVVIAGTVFVAYHRDSEYKNTKMLQPDYDSTHSTRTVTVKQDQHSR 236
Qy 228 GWFVGSYRRRTTVNCELMDMFARSNPPDFVATGDTVMSPFSGEDDHENKHEKWP 287
Db 237 GSTW-LYRETCNLNCMTVITITARSKYPHFFATSTGVDVDSIPYNGTNRNASYFGENAD 295
Qy 288 FVSVINNYKVVDYQNRGTVPGLKTRI--FLDREYTLSE--KHLKNMSYCPPLTLWKAFYN 344
Db 296 KFFIFPNYTVSDGFPNSALETHRLVAFLEADSVISWDIQDEKNT-CQLTFWEASER 354
Qy 345 GIOTHSYGVHVFANDITASFTTSKEDMKFNTYHCLNBEIKAEIEKKY-AKVNSTHSHK 403
Db 355 TIRSEADSYHFSKATATATFLSKQEVNMSDCLDCVRDEAINKLOQIFNTSYNTYK 414
Qy 404 YGDLKYFKTDGGLYLVWQPLIQNRLDANKLNNETYSRRSRQAESTTDPMMEMTNGA 463
Db 415 YGNVSVFETTGGLVVFQGIKQSLVELERLANRSSLNLTHNTRKST-----DGNNA 467
Qy 464 G--GEYSSENSITVAQVQYAYDNIRIRINNILEDLSKAWCREQHRALVWNELSKINPTS 521
Db 468 THLSNMESVHNLVYAQLQFTYDTRLGYINRALAQIAEAWCVDQRTLEVFKELSKINPSA 527
Qy 522 VMSMIYNNRPVSARKIGDVISVSNICIVVDOTSVSLHLSLRLLSASDEKCFRPPVTFKPMN 581
Db 528 ILSAIYNKPIAARFMDGLVGLASCVTINQTSVKVLRDMN-VKESPGRCYSRPVVIENFAN 586
Qy 582 DSTIYKQGLGVNNEILLTTTLYLETCQENTYFYFOAKTDMYIKNYEHLKTVPLSSITLD 641
Db 587 SSVYQYQGLGEDNELLGNHTECQLPSLKIFAGNSAYEYVDYLFKRMIDLSISTVD 646
Qy 642 TFIALNFTLENDVFKVIELYTRDEKRLSNVFDIETMFRYNYAQRVSGRLKDLDLST 701
Db 647 SMIALDIPLENTDFRVELYSQKELRSSNVDFLEIMREFNSYKQRYVEDKVDV-- 703
Qy 702 NRNOFVDAFGLMDLCAVGQTVNVNAGVATLFPSSIVTGTFINFKNPFGLMLIIVIG 761
Db 704 PLPPYLKGLDMLGSLGAAGKAVGAIGAVGAVASVVEGATFLKPNPFGAFTIILVAIA 763
Qy 762 VLFAIYELTKTKIYETAPIKMIYPEI-----DKLKE-----R 794
Db 764 VVIILYLYTRQRLCQPLQNLPLVLSADGTTVTSGNTKOTSLQAPPSEYSEVNSGR 823
Qy 795 EG-----KSETAPISELELERIVLAM-----HIHQONS-----HMETKTR-KDPKOS 835
Db 824 KGPGPSDDASTAPPYTNQAYQMLLALVRLDAEQRAQNGTDSLQDGTQDKGQKPN 883
Qy 836 ILTRAQNLKRSYGNLKNASVE 860
Db 884 LLDLRLH---RKNGYRHLKSDSEE 905

RESULT 29

US-08-804-439A-19
; Sequence 19, Application US/08804439A
; Patent No. 6015565
; GENERAL INFORMATION:
; APPLICANT: Rose, Timothy M.
; APPLICANT: Bosch, Marnix L.
; APPLICANT: Strand, Kurt
; TITLE OF INVENTION: GYCOPROTEIN B OF THE RRVH/KSHV
; STREET: 4225 Executive Square, Ste 1400
; CITY: La Jolla
; STATE: CA
; COUNTRY: USA
; NUMBER OF SEQUENCES: 113
; CORRESPONDENCE ADDRESS:
; ADDRESS: Fish & Richardson P.C.
; STREET: 4225 Executive Square, Ste 1400
; CITY: La Jolla
; STATE: CA
; COUNTRY: USA

ZIP: 92037
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/804,439A
FILING DATE: February 21, 1997
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Haile, Lisa A.
REGISTRATION NUMBER: 38,347
REFERENCE/DOCKET NUMBER: 09176/004001
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 678-5070
TELEFAX: (619) 678-5099
TELEX:
INFORMATION FOR SEQ ID NO: 19:
SEQUENCE CHARACTERISTICS:
LENGTH: 907 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-804-439A-19
Query Match 24.8%; Score 1128; DB 3; Length 907;
Best Local Similarity 30.9%; Pred. No. 3,9e-79;
Matches 286; Conservative 184; Mismatches 353; Indels 102; Gaps 25;
Qy 14 LWLYOVALYSLSIAETG--VTSPNTATWSTESPLTGHVGHDS--HGERGNNEIRDS 69
Db 5 IWCL--VVCUNLCIVCLGAASSTSTRTGTSATHS-----HHSHTTAAHSRSSVSQRTV 58
Qy 70 EQ-----NKNYIG-----SPSTPPYRVCASGVGDVFRFOTDHC-----PDA 108
Db 59 SSTQVSHGVNETIYNTTLTKYGVGVVNTTKYPYRVCMAQGTDLIRFERNIVCTSMKPIN 118
Qy 109 SDMVHSEGLLIYKONIIPFMRVRYKRVKVTSTVYNGIYSDSITNOHTFFYKSIPEWET 168
Db 119 EDL--DEGIMVVKRNIVAHFTKVRVQKVLTRRSVAYIHTTYLLGSNTSEYVAPPWMEI 176
Qy 169 EKMDTIYOCNSLRNLTTGNNLLTVVDRDDI--NMTVFLQPVGDVTPDVKRYGSOPELYLEP 227
Db 177 HHINSHQCYSSYSSRVVIAGTVFVAYHRDSEYKNTKMLQPDYDSTHSTRTVTVKQDQHSR 236
Qy 228 GWFVGSYRRRTTVNCELMDMFARSNPPDFVATGDTVMSPFSGEDDHENKHEKWP 287
Db 237 GSTW-LYRETCNLNCMTVITITARSKYPHFFATSTGVDVDSIPYNGTNRNASYFGENAD 295
Qy 288 FVSVINNYKVVDYQNRGTVPGLKTRI--FLDREYTLSE--KHLKNMSYCPPLTLWKAFYN 344
Db 296 KFFIFPNYTVSDGFPNSALETHRLVAFLEADSVISWDIQDEKNT-CQLTFWEASER 354
Qy 345 GIOTHSYGVHVFANDITASFTTSKEDMKFNTYHCLNBEIKAEIEKKY-AKVNSTHSHK 403
Db 355 TIRSEADSYHFSKATATATFLSKQEVNMSDCLDCVRDEAINKLOQIFNTSYNTYK 414
Qy 404 YGDLKYFKTDGGLYLVWQPLIQNRLDANKLNNETYSRRSRQAESTTDPMMEMTNGA 463
Db 415 YGNVSVFETTGGLVVFQGIKQSLVELERLANRSSLNLTHNTRKST-----DGNNA 467
Qy 464 G--GEYSSENSITVAQVQYAYDNIRIRINNILEDLSKAWCREQHRALVWNELSKINPTS 521
Db 468 THLSNMESVHNLVYAQLQFTYDTRLGYINRALAQIAEAWCVDQRTLEVFKELSKINPSA 527
Qy 522 VMSMIYNNRPVSARKIGDVISVSNICIVVDOTSVSLHLSLRLLSASDEKCFRPPVTFKPMN 581
Db 528 ILSAIYNKPIAARFMDGLVGLASCVTINQTSVKVLRDMN-VKESPGRCYSRPVVIENFAN 586
Qy 582 DSTIYKQGLGVNNEILLTTTLYLETCQENTYFYFOAKTDMYIKNYEHLKTVPLSSITLD 641


```

;
;   NUMBER OF SEQUENCES: 10
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE: Howson and Howson
;   STREET: Spring House Corporate Center, P.O. Box 457
;   CITY: Spring House
;   STATE: Pennsylvania
;   COUNTRY: USA
;   ZIP: 19477
;
;   COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: Patent in Release #1.0, Version #1.30
;
;   CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US 60/015,717
;   FILING DATE: 23-APR-1996
;   ATTORNEY/AGENT INFORMATION:
;   NAME: Kodtsoff, Cathy A.
;   REGISTRATION NUMBER: 33,980
;   REFERENCE/DOCKET NUMBER: WST666APT
;
;   TELECOMMUNICATION INFORMATION:
;   TELEPHONE: 215-540-9200
;   TELEFAX: 215-540-5818
;
;   INFORMATION FOR SEQ ID NO: 2:
;   SEQUENCE CHARACTERISTICS:
;     LENGTH: 907 amino acids
;     TYPE: amino acid
;     TOPOLOGY: linear
;
;   MOLECULE TYPE: protein
;
;   SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-171-699-2

Query Match      24.8%; Score 1128; DB 4; Length 907;
Best Local Similarity 30.9%; Pred. No. 3.9e-79;
Matches 286; Conservative 184; Mismatches 353; Indels 102; Gaps 25;

Qy 14 LWLYQVALSLSAETG--VTSPNTATSTESPLTGHGYTHDSS--HGERGNENRDS 69
Db 5 IWCL--VVCVNLICVLGAIVSSSSSTGTSATHS----HSSHTTSAHSRSGVSQRVT 58
Qy 70 EQ-----NKNIVG-----SPSTFPYRVCSASGVGVFRFQTDHVC-----PDA 108
Db 59 SSQTVSHGVNETIYNTTLKYGDVVGVTNTKYPYRVCSMAQGTDLIRPERNIVCTSMKPIN 118
Qy 109 SDMVHSGILLIYKQNIIPMFVRKYRKVVTSTVYNGIYSDSITNQHTFYKSIPEWET 168
Db 119 EDL--DEGIMVVKRNIYVAHTFKRVYQKVLTFRRSYAYIHTYLLGSNTTEYVAPPMEI 176
Qy 169 EKMDTIYQCFNSLRNLGNGLLTYVDRDDI--NMTVFLQPVGVTPDVKRYGSPELYLEP 227
Db 177 HHINSHQCVSSYSRVVAGTVFVAYHSDSYENKTMQLMPDDYSNTHSTRVYVTKDQHSR 236
Qy 228 GWFVGSYRRRTTNCVCLMDMFARNPPDFVTATGTVEMSPFWSGEDDHENKMEKPW 287
Db 237 GSTW-LYRETCLNLCMVTITARSKYPHFPFATSTGWDVSDISPFYNGTNRNNSYFGENAD 295
Qy 288 FVSVINNYKVVDYQNRGTVPVGLKTRI--FLDREYTLISWE-KHLKMSYCPPLTKAFYN 344
Db 296 KFFIFPNYIVSDGFRNPSALETHRLVAFLEADSVISWDIQEKNVT-COLTFWEASER 354
Qy 345 GIQTEHSGSYHFVANDITASFTTSKEDMKBFNTYTHCLNEEIKAEIEKKY-AKVNSTHSHK 403
Db 355 TIRSEADSYPHFSAKMTATFLSKQEVNMSDSALDCVRDEAKNKLOQIENSTVNTQYEK 414
Qy 404 YGDLKYPTDGGYLVVQPIQNLRLDLAKNKLNNETVRRSRQAESTTDPMMEMTGNGA 463
Db 415 YGNVSVPFETTGGLVVFQGIKQKSLVELERLANRSSLNLTNHRNTRKST-----DGNNA 467

```

RESULT 32

```

PCT-US94-04180-2
; Sequence 2, Application PC/TUS94040180
; GENERAL INFORMATION:
; APPLICANT: Wistar Institute of Anatomy, Biology
; APPLICANT: Government of USA, Dept.
; APPLICANT: Health and Human, Services
; TITLE OF INVENTION: Recombinant Cytomegalovirus Vaccine
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Howson and Howson
; STREET: Spring House Corporate Center, P.O. Box 457
; CITY: Spring House
; STATE: Pennsylvania
; COUNTRY: USA
; ZIP: 19477
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US94/04180
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/048,978
; FILING DATE: 16-APR-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Bak, Mary E.
; REGISTRATION NUMBER: 31,215
; REFERENCE/DOCKET NUMBER: WST66CPT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 215-540-9200
; TELEFAX: 215-540-5818
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
;   LENGTH: 907 amino acids
;   TYPE: amino acid
;   TOPOLOGY: linear

```

```

Qy 464 G--GEYSSSENSITVAQVAYDNLRIRINNILEDLSKAWCREOHRALVWNLKSNKPTS 521
Db 468 THLSNMESVHNLVYAQLQFTYDTRLGYINRALAQIAEAMCVDRRTLEVFKELSKINPSA 527
Qy 522 VMSMIYNRPVSAKRIGDIVSVSNICIVVDQTSVSLHSLRLLSASDEKCFSPRPVTFKPMN 581
Db 528 ILSALIYNKPFAARFMGDLGLASCVTINQTSVKVLRDMN-VKESPCRCYSRPPVIFNFAN 586
Qy 582 DSTIYKQGLGVNNEILLTYYLTCTCOENTEYFOAKTDMYIKYNYEHLKTVPLSSITLTD 641
Db 587 SSVYQVQGLGEDNEILLGNHRTTEECQLPSLKIFIAGNSAYEYVDYLFKRMIDLSSISTVD 646
Qy 642 TETALNFTLLENVDVDFKVIETYDEKRLSNVPDIETMEREYNYAQRVSLRKLDDLDLST 701
Db 647 SMIALDIDPLENTDFRVELYSQKELRSSNVFDEEIMREFNSYKORVKVVEKVDV--- 703
Qy 702 NRNQFVDAFGSLMDDLGAQGVTVYNAVSGVATLTFSSIVTGFNFINKPFGGMLMIIVIG 761
Db 704 PLPPYLKGLDGLMSGLGAAGKAVGAIVGAVGVAVSVVEGVATFLKNPFGAFTIILVAIA 763
Qy 762 VLFATYFLTKTKIYETAPIKMIYPEI-----DKLKE-----R 794
Db 764 WIIIIYIYTRQRLCMQPLNLFPYLVSGADGTTVTSGNTKDTSLQAPPSESVYNSGR 823
Qy 795 EG-----KSEIAPISEEELERIVLAM-----HIHQNS-----HMETKTR-KDPKDS 835
Db 824 KGGPSSSDASTAAPPTNEQAYQMLLALVRLDABEORAQNGTDSLDDGTGTDKQKQPN 883
Qy 836 ILTRAQNMRLKRSYGNLKNNAESVE 860
Db 884 LLDRLRH---RKNYVRLKXSDDEE 905

```

MOLECULE TYPE: protein
PCT-US94-04180-2

```

Query Match          24.8%; Score 1128; DB 5; Length 907;
Best Local Similarity 30.9%; Pred. No. 3.9e-79;
Matches 286; Conservative 184; Mismatches 353; Indels 102; Gaps 25;

QY 14 LWLYQVALYSLSIAETG--VTPSPNTATWSTESPLTGHYTHDSS--HGERGNNENRDS 69
DB 1 WCL--VVCVNLCLVCLGAAVSSSTRGTSATHS---HSSHTTSAHRSRSGSVQRVT 58
QY 70 EQ-----NKNIY-----SPSTFPYRVCSASGVGDVRFOTDHYC-----PDA 108
DB 59 SSQTVSHGNETIYNTTLKYGVGVGVNTTKYPRVCSMAQGTDLIRFERNICTSMKPIN 118
QY 109 SDMVHSEGLILLYKONIIPFMRKYRKVVTSTTVYNGIYSDSIHQHTFYKSIETPWT 168
DB 119 EDL--DEGLIMVYKRNIVAHTEFKVYQKVLTPRSYAVIHTTYLLGNSNTEYVAPMWEI 176
QY 169 EKMDTIYQCFNSRLNLTGNNLLTYVDRDDI--NMTVFLQPVGDVTPDKVRYGSQPELYLEP 227
DB 177 HHNSHSQCYSSYSRVIAGTVFVAYHRDSYENKTMQAMPDDYDSNTHSTRYTVTKDQWHSR 236
QY 228 GWFWSYRRRTTVNCELMDMPARSPPDPFVATGDTVEMSPFWSGEDDHENKMKHPW 287
DB 237 GSTW-LYRETCLNLCMTITITARSKYPYHFFATSTGDDVVDISPFYNGTNRNASYFGENAD 295
QY 288 FVSVINNYKVQYQNGTVPLCKTRI--FLDREYTLSE--KHLKNMSYCPCLTLKAFYN 344
DB 296 KFFIFPNYTVISDFGRPSALETHRLVAFLEADSVISWDIODEKNVT-CQLTFWEASER 354
QY 345 QIOTHSYHGFVANDITASFSTSKEDMKEFNVTYHCLNEEIKAEIKKY-AKVNSTHSH 403
DB 355 TIRSEADSYHPSAKMTATFLSKQEVNMSALDCYRDEAINKLOIFNTSYNQTEYK 414
QY 404 YDLYKFTDGLYLWVQPLIQNRLLDANKLNNTYRRSRQAEASTDDPMEMTNGA 463
DB 415 YGNVSVEFTTGLVFWQGIQKSLVELERLANRSSLNLTNRTRKST-----DGNA 467
QY 464 G--GEYSSENSITVAQVAYDNLIRINNILEDLSKACREOHRALVWHELKINTS 521
DB 468 THLSNMESVHNLYVLAQLOFTYDLRGYNRLAAQIAEAWCVDQRRTEVFEKLSKINPSA 527
QY 522 VMSMIYRNPVSRAKRGDVISVNCIIVDQTSVSLHSLRLLSASDEKCFSPRPVTFKPMN 581
DB 528 ILSAIYNKPIARFWDVLGLASCTVINGTSVKVLRDMN-VKESPCRCYSRPFVIFNAN 586
QY 582 DSTIYKGLGVNNEILLTTLTYLETQENTYFYQAKTDMYIYKNYEHKLTVPPLSSITLD 641
DB 587 SSVYQVQGLGEDNEILLGHRTEECPQLSKLFIAGNSAYEYVDYLFKRMIDLSISITVD 646
QY 642 TFIALNFTLLENVDPKVIETLYTRDEKRLSNVFDIETMFREYNYAQRVSGLRKOLLDSLT 701
DB 647 SMIALDIPLENTDFRVLLEYLSQKELRSNVFDLEIMREFNSYKQRYVEDKVD---703
QY 702 NRNQVDFAGSLMDLGLAGVGVTVNAVSGVATLFSSIVTGFINFINKNPFGLMLIIVIG 761
DB 704 PLPPYKGLDMLGSLGAGKAVGAUGVAGVASVEGVATFLKNPFGAFTIILVAIA 763
QY 762 VLFAIYFLTKTKIYETAPIKMIYPEI-----DKLKE-----R 794
DB 764 VVIIIVLYTRQRLCMQPLQNLFPYLVASDGTTSNGTKDTSLOAPPSYEESVYNSGR 823
QY 795 EG-----KSEIAPISSEELERIVLAM-----HIHQONS-----HMETKTR-KDPKDS 835
DB 824 KGPFPSSDASTAAPPTNEQAYOMLLALVRLDAEQRAQQNGTSLDQGTGTQDQKQPN 883
QY 836 ILTRAQNMRLKBSGYSNLKNASVE 860
DB 884 LLDRLR---RKNGRHLKSDSEEE 905

```

RESULT 33

```

US-08-220-151-9
; Sequence 9, Application US/08220151
; Patent No. 5529780
; GENERAL INFORMATION:
; APPLICANT: Paoletti, Enzo
; TITLE OF INVENTION: NUCLEOTIDE AND AMINO ACID SEQUENCES OF
; TITLE OF INVENTION: CANINE HERPESVIRUS GB, GC AND GD AND USES THEREFOR
; NUMBER OF SEQUENCES: 91
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Curtis, Morris & Safford
; STREET: 530 Fifth Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA: US/08/220,151
; APPLICATION NUMBER: US/08/220,151
; FILING DATE: 30-MAR-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Frommer, William S.
; REGISTRATION NUMBER: 25,506
; REFERENCE/DOCKET NUMBER: 454310-2540
; TELEPHONE: (212) 840-0712
; TELEFAX: (212) 840-0712
; TELEX: 425066 CURTMS
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 906 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FRAGMENT TYPE: N-terminal
; US-08-220-151-9

```

```

Query Match          24.7%; Score 1121.5; DB 1; Length 906;
Best Local Similarity 30.4%; Pred. No. 1.2e-78;
Matches 282; Conservative 186; Mismatches 350; Indels 111; Gaps 25;

QY 14 LWLYQVALYSLSIAETGVTSPNTATWSTESPLTGHYTHDSSHGERGNN-ENRD--- 68
DB 5 WCL--VVCVNLCLVCLGAAVSSSSTSHATSS-----THNGSHTSRSTTSAQTRSVYSQ 55
QY 69 -----SEQNKNIY-----SPSTFPYRVCSASGVGDVRFOTDHYC----- 105
DB 56 HVTSEAVSHRANETIYNTTLKYGVGVGVNTTKYPRVCSMAQGTDLIRFERNICTSMK 115
QY 106 PDASDMVHSGILLIYKONIIPFMRKYRKVVTSTTVYNGIYSDSIHQHTFYKSIETP 165
DB 116 PINEDL--DEGLIMVYKRNIVAHTEFKVYQKVLTPRSYAVIHTTYLLGNSNTEYVAPM 173
QY 166 WETEKMDTIYQCFNSRLNLTGNNLLTYVDRDDI--NMTVFLQPVGDVTPDKVRYGSQPELY 224
DB 174 WEIHHINKFAQCYSSYSRVIAGTVFVAYHRDSYENKTMQAMPDDYDSNTHSTRYTVTKDQW 233
QY 225 LEPQWFGSVYRRRTTVNCELMDMPARSPPDPFVATGDTVEMSPFWSGEDDHENKMKHE 284
DB 234 HSRGSTW-LYRETCLNLCMTITITARSKYPYHFFATSTGDDVVDISPFYNGTNRNASYFGE 292
QY 285 KPWFPVSVINNYKVY-DYQNGRTVP-LGKTRIFLDREYTLSE--KHLKNMSYCPCLTLKWA 341
DB 293 NADKFFIFPNYTVISDFGRPSALETHRLVAFLEADSVISWDIODEKNVT-CQLTFWEA 351
QY 342 FYNGIOTHSYHGFVANDITASFSTSKEDMKEFNVTYHCLNEEIKAEIKKY-AKVNST 400

```

```

Db 352 SERTIRSEADSYHFSSAKMTATFLSKQEVNMSDLSALDCVRDEAINKLQOIFNTSYNOT 411
Qy 401 HSKYGLDKYFXTDGLGLVWQPLQNLRLDADKNNETYS---RRSRQAESTTDPME 457
Db 412 YEKYGNVSVFTSGLVFWQGIKQKSLVELELRANRSSLNITHTRRSTSDNNTTLLSS 471
Qy 458 MTGNAGGEYSSENSITVAQVQYADNLRIRINNILEDLSKAWCREQHRALVWNLSKI 517
Db 472 M-----ESVHNLVYAQLQFTYDTLRGYINRALAQIAEAWCVDQRRITLVEFKELSKI 522
Qy 518 NPTVMSMIYNRPVSAKRIGDVISVNSCIVVDQTSVSLHKSRLLSASDEKCFSPVPVF 577
Db 523 NPSAILSAIYNKPIAARFNGDVLGLASCVTINQTSVKVLRDMN-VKESPCRCYSRPVWIF 581
Qy 578 KFMNDSTIYKQGVNNEILLTITTYLETQENTYFQAKTDMYIYKNEYHLKTVPLSSI 637
Db 582 NFANSSVYQVQGLGDENEILLGNHRTBECQPSLKIFAGNSAYEYVDYLFKRMIDLSSI 641
Qy 638 TTLDTFIALNFTLLENVDFKVIELYTRDEKRLSNVFDIETMFRYNYVAQRVGLRKDLL 697
Db 642 STVDSMIALDIDPLENTDFRVELYSQKELRSSNVFDEEIMREFNSYKQRYKVEDKV 701
Qy 698 DLSTNRNQVDAFGSLMDLGAQGVTVVNAVSGVATLFSIVTGFINFKNPFGMLMII 757
Db 702 D---PLPPYLKGLDLDMSGLGAAGKAVGAIGAVGAVASVVEGVATFLKNPFGAFTIL 758
Qy 758 VVIGVLFALYFLTKTKIYETAPIKMIYPEI-----DKLKE----- 793
Db 759 VAIAVVIITYLYTRQRLCTQPLQNLFFPYLVLSADGTTVTGSGTKTSLQAPPSYBESVY 818
Qy 794 ---REG-----KSEIAPISEEELERIVLAM-----HIHQONS-----HMETKTR-KD 831
Db 819 NSGRKGCPSSDASTAAPPTNEQAYQMLALARLDAEQRAQNGTSLDGTGTQDKG 878
Qy 832 PKOSILTRAQNMRLKRGSYNLKNAESVE 860
Db 879 QKPNLLDLRLH---RKNGYRLKXDSDEE 904

RESULT 34
US-08-413-118-9
; Sequence 9, Application US/08413118
; Patent No. 5688920
; GENERAL INFORMATION:
; APPLICANT: PAOLETTI, ENZO
; APPLICANT: LIMBACH, KEITH J.
; TITLE OF INVENTION: NUCLEOTIDE AND AMINO ACID SEQUENCES OF
; TITLE OF INVENTION: CANINE HERPESVIRUS GB, GC, AND GD AND USES THEREFOR
; NUMBER OF SEQUENCES: 128
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CURTIS, MORRIS & SAFFORD, P.C.
; STREET: 530 FIFTH AVENUE, 25TH FLOOR
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: UNITED STATES OF AMERICA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/413,118
; FILING DATE: 29-MAR-1995
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/220,151
; FILING DATE: 30-MAR-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: FROMMER, WILLIAM S.
; REGISTRATION NUMBER: 25,506
; REFERENCE/DOCKET NUMBER: 454310-2670
; TELECOMMUNICATION INFORMATION:

```

```

; TELEPHONE: (212) 840-3333
; TELEFAX: (212) 840-0712
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 906 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FRAGMENT TYPE: N-terminal
; US-08-413-118-9

Query Match 24.7%; Score 1121.5; DB 1; Length 906;
Best Local Similarity 30.4%; Pred. No. 1.2e-78;
Matches 282; Conservative 186; Mismatches 350; Indels 111; Gaps 25;

Qy 14 LMVLYQVALYSLSIAETGVTSPPTATWSTESPLTGHYTHDSSHGGRNN-ENRD--- 68
Db 5 IWCL--VVCNLICIVCLGAAVSSSSSTSHATSS-----THNGSHTSRITSAQTRSVYSQ 55
Qy 69 -----SEEQNKIY-----SPSTFPYRVCSASGVGDVFRFQTDHVC----- 105
Db 56 HVTSEAVSHRANETIYNTTLKYGDVVGVTNTTKYRVVCSMAQGTDLIRFERNICTSMK 115
Qy 106 PDASDMVHSEGIILYKQNIIPMPFRVRKYRVTSTTVYNGIYSDSITNQHTFYKSIIEP 165
Db 116 PINEDJ--DEGMVYKRNIVAHFKRVYQKVLTFRRSYAYIYTYLLGNSNTEYVAPPM 173
Qy 166 WETEKMDTIYQCNSLURLTGNLLTYVDRDDI--NMTVFLQPVGDVTPDVKRYGSOPELY 224
Db 174 WEIHHINKPAQCYSYRVIGGTVPVAYHRDSYENKTMQLIPDDYNTHSTRVYTVKQDW 233
Qy 225 LEPCFWGYSRRRTTVNCELMDFARSNPPFPFVATGDTVEMSPFWGDEDDHENKMH 284
Db 234 HSRGSTW-LYRETCLNCLMTTITARSKYPIHFATSTGDDVYVISPFGYNGTNRNASYFGE 292
Qy 285 KPWFSVNNYKVV-DYQNGRTVP-LGKTRI FLDRREYTLSE--KHLKMSYCPLTLWKA 341
Db 293 NADKFFIFPNYIIVSDFGRPNAPETHRLVAFLEADSVISWDIQDEKNT-CQLTFWEA 351
Qy 342 FYNGIQTEHSGSYHFVANDITAFSTTSKEDMEFNTYTHCLNBEIEKAEIEKKY-AKVNST 400
Db 352 SERTIRSEADSYHFSSAKMTATFLSKQEVNMSDLSALDCVRDEAINKLQOIFNTSYNOT 411
Qy 401 HSKYGLDKYFXTDGLGLVWQPLQNLRLDADKNNETYS---RRSRQAESTTDPME 457
Db 412 YEKYGNVSVFTSGLVFWQGIKQKSLVELELRANRSSLNITHTRRSTSDNNTTLLSS 471
Qy 458 MTGNAGGEYSSENSITVAQVQYADNLRIRINNILEDLSKAWCREQHRALVWNLSKI 517
Db 472 M-----ESVHNLVYAQLQFTYDTLRGYINRALAQIAEAWCVDQRRITLVEFKELSKI 522
Qy 518 NPTVMSMIYNRPVSAKRIGDVISVNSCIVVDQTSVSLHKSRLLSASDEKCFSPVPVF 577
Db 523 NPSAILSAIYNKPIAARFNGDVLGLASCVTINQTSVKVLRDMN-VKESPCRCYSRPVWIF 581
Qy 578 KFMNDSTIYKQGVNNEILLTITTYLETQENTYFQAKTDMYIYKNEYHLKTVPLSSI 637
Db 582 NFANSSVYQVQGLGDENEILLGNHRTBECQPSLKIFAGNSAYEYVDYLFKRMIDLSSI 641
Qy 638 TTLDTFIALNFTLLENVDFKVIELYTRDEKRLSNVFDIETMFRYNYVAQRVGLRKDLL 697
Db 642 STVDSMIALDIDPLENTDFRVELYSQKELRSSNVFDEEIMREFNSYKQRYKVEDKV 701
Qy 698 DLSTNRNQVDAFGSLMDLGAQGVTVVNAVSGVATLFSIVTGFINFKNPFGMLMII 757
Db 702 D---PLPPYLKGLDLDMSGLGAAGKAVGAIGAVGAVASVVEGVATFLKNPFGAFTIL 758
Qy 758 VVIGVLFALYFLTKTKIYETAPIKMIYPEI-----DKLKE----- 793
Db 759 VAIAVVIITYLYTRQRLCTQPLQNLFFPYLVLSADGTTVTGSGTKTSLQAPPSYBESVY 818
Qy 794 ---REG-----KSEIAPISEEELERIVLAM-----HIHQONS-----HMETKTR-KD 831

```

Db 819 NSGRKGPSSDASTAAPPYNEQAYQMLLALRLDAEQRAQNGTDSLDCGTQDKG 878
Qy 832 PKDSILTRAQNMRLKRSGYSLNKAESVE 860
Db 879 QKPNLLDLRLH---RKNGYRHLKDSDEE 904

RESULT 35
US-08-473-446-9
; Sequence 9, Application US/08473446
; Patent No. 6017542
; GENERAL INFORMATION:
; APPLICANT: PAOLETTI, ENZO
; APPLICANT: LIMBACH, KEITH J.
; TITLE OF INVENTION: NUCLEOTIDE AND AMINO ACID SEQUENCES OF
; TITLE OF INVENTION: CANINE HERPESVIRUS gb, gc, and gd AND USES THEREFOR
; NUMBER OF SEQUENCES: 128
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CURTIS, MORRIS & SAFFORD, P.C.
; STREET: 530 FIFTH AVENUE, 25TH FLOOR
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: UNITED STATES OF AMERICA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/473,446
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/413,118
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: FROMMER, WILLIAM S.
; REGISTRATION NUMBER: 25,506
; REFERENCE/DOCKET NUMBER: 454310-2670
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 840-3333
; TELEFAX: (212) 840-0712
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 906 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FRAGMENT TYPE: N-terminal
US-08-473-446-9

Query Match 24.7%; Score 1121.5; DB 3; Length 906;
Best Local Similarity 30.4%; Pred. No. 1.2e-78;
Matches 282; Conservative 186; Mismatches 350; Indels 111; Gaps 25;

Qy 14 LWLYQVALYSLSIAETGTPSPNTATWSTSPLTGCHYGTDDSHGEGNN-ENRD--- 68
Db 5 IWL--VVCNLICVCLGAAVSSSSSTSHATSS-----THNGSHTSRTTSAQTRSVYSQ 55
Qy 69 -----SEONKNIYG-----SPSTFPYRVCASGVGDVRFQTDHVC----- 105
Db 56 HVTSEAVSHRANETIYNTTKYGDVGVNTTKYPIRVCMAQGTDLIRFERNICTSMK 115
Qy 106 PDASDMVHSEGILLYKQNIIPFMRVRKRVKVTSTTYNGIYSDSITNOHTFYKSTEP 165
Db 116 PINEDL--DEGLMVVYKKNIVAHTEKRVYQKVLTFRRSYAYIYTYLLGNTGVAPPM 173
Qy 166 WETEKMDTYQCFNSLRNLNTGNLLTYVDKDI--NMTVFLOPVDGVTPDKVRYGSGQELY 224
Db 174 WEIHHINKFAQCYSSYSRVIGGVTFVAVHRDSYENKTMQLIPDDYSNTHSTRYVTVKDQW 233

Qy 225 LEPGWFSGYRRRTTVNCELMDFAFNSPPDFVFVATGDTVEMSPFWSGDDHDKMHE 284
Db 234 HSRGSLW-LYRETCLNLCMLTITARSKYPHFATSTGDDVYISPFYNGTRNASYGE 292
Qy 285 KPWFSVSNVYKVV-DYONGTVP-LGKTRIFLDREYTLSE-KHLKNMSYCPJTLWKA 341
Db 293 NADKFFIPPNYTVISDFGRPNAAPETHRLVAFLEADSVISWDIQDEKNVT-CQLTFWEA 351
Qy 342 FYNGIOTEGSGYHFVANDITASFTTSKEDMKENFTYHCLNEEIKAEIEKKY-AKNST 400
Db 352 SERTIRSEAEADSYHFSAKMTATFLSKQEVNMSDALSALDCVRDEAINKLOQIFNTSYNQ 411
Qy 401 HSKYGDLYKFKTDGGLYLVQPLIQNRLDLDAKNLNNETYS---RRSRQAEESTDPNME 457
Db 412 YEKYGNVSFVETSGGLVVFQGIQKSLVELERLANRSSLNITHTRRSISDNWTHLSS 471
Qy 458 MTGNAGGEYSSENSITVAQVQYAYDNLIRINILEDLSKAWCREQHRAALVNELSKI 517
Db 472 M-----ESVHNLVYAQLQFTYDTLRGVINRALAQIAEAMCVDQRRTLVFKELSKI 522
Qy 518 NPTSVMSMIYNRPVSAKRIGDIVSNVICVDDOTSVSLHKSLLRLLSASDEKCFRPPVTF 577
Db 523 NPSAILSAIYNKPIAARFMGDLGLASCVTINQTSVKVLRDMN-VKESPGRCYSRPVIF 581
Qy 578 KFMNDSTIYGQLGQVNNELLTTTLETQENETVYFOAKTDMYIYKNEYHLKTVPLSSI 637
Db 582 NFANSSVVOYQGLGEDNEILLGNHRTBECQPLSKIFIAGNSAYEYDYLFRMIDLSSI 641
Qy 638 TLTDTFIALNFTLLENVDKVIELYTRDEKRLSNVFDIETMFRYNYVAORVSGLRKDLL 697
Db 642 STVDSMIALDIDPLENTDFRLELYSQELRSSNVFDEEIMREBFSYKQRYKVEDKV 701
Qy 698 DLSTNRNQFVDAFGSLMDDLGAQCQTVNAVSGVATLFSSTVGTGFIKPKFGMLMI 757
Db 702 D---PLPPYLKGLDGLMSGLGAAGKAVGAIVGAVASVVEGVATFLKNPFGAFTIL 758
Qy 758 VWIGVFAIYFLTKTKIYETAPIKMIYPEI-----DKLKE----- 793
Db 759 VAIAVIITYLIYTRQRLCTQPLNQNLFPYLVSDGTTVTSGSTKDTLSLQAPPSYEEVY 818
Qy 794 ---REG-----KSEIAPISEEELERIVLAM-----HIHQONS-----HMETKTR-KD 831
Db 819 NSGRKGPSSDASTAAPPYNEQAYQMLLALRLDAEQRAQNGTDSLDCGTQDKG 878
Qy 832 PKDSILTRAQNMRLKRSGYSLNKAESVE 860
Db 879 QKPNLLDLRLH---RKNGYRHLKDSDEE 904

RESULT 36
US-08-804-439A-20
; Sequence 20, Application US/08804439A
; Patent No. 6015565
; GENERAL INFORMATION:
; APPLICANT: Rose, Timothy M.
; APPLICANT: Bosch, Marnix L.
; APPLICANT: Strand, Kurt
; TITLE OF INVENTION: GLYCOPROTEIN B OF THE RPHV/KSHV
; TITLE OF INVENTION: SUBFAMILY OF HERPES VIRUSES
; NUMBER OF SEQUENCES: 113
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 4225 Executive Square, Ste 1400
; CITY: La Jolla
; STATE: CA
; COUNTRY: USA
; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30

```

: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/804.439A
: FILING DATE: February 21, 1997
: CLASSIFICATION: 424
: ATTORNEY/AGENT INFORMATION:
: NAME: Haile, Lisa A.
: REGISTRATION NUMBER: 38,347
: REFERENCE/DOCKET NUMBER: 09176/004001
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (619) 678-5070
: TELEFAX: (619) 678-5099
: TELEX:
: INFORMATION FOR SEQ ID NO: 20:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 830 amino acids
: TYPE: amino acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: US-08-804-439A-20

Query Match      23.1%; Score 1050.5; DB 3; Length 830;
Best Local Similarity 32.3%; Pred. No. 3.6e-73;
Matches 244; Conservative 160; Mismatches 293; Indels 59; Gaps 15;

Qy 83 FPRVCSAGVGDFRFQTDHVC--PDASDMVHSEGLLIYKQNIIPMFVRVRYKRVKVT 141
Db 36 YPRFICSIAGTDLMRDRDISCPYKSNAMSGFFIIYKTIETIYTPVRYKKELT 95
Qy 142 STVYNGIYSDSITNQHTFYKSIETPEWETEKMDTIYQCFNSRLN--TGGNLLTYVDRDIN 199
Db 96 QSSYRDVGVVYFLDRVTMGLAMPVYEANLVNSHAQCYSAVAMKRPDGTVFSAFHEDNNKN 155
Qy 200 MTVPLOPVDGVTDPVKRYGSOPELYLEPGFWGYSRRRTTVNCCLMDMPARSPPDPFFV 259
Db 156 NTNLNLPFLNFKSITNKRFTITKEPYFARGPLW--LYSTSTSLNCIVTEATKAKYPSFYA 214
Qy 260 TATGDTVMSPFWGSD--DHENKMKHEKPMFVSVINNYKV--DYON--RGTVPLGKTRIFL 315
Db 215 LTTGEIVGEGPFFNGSKGHFAEPLEK--LTLLENVTMIEDLMNGMNGATTILVRKIAFL 271
Qy 316 DREYTLSEKHLKNNMSYCPDLTLWKAFYNGIQTEHSGSYHFVANDITASTTTSKEDMKEF 375
Db 272 EKADTLFSWEIKEENESVCMKHWTTVTHGLRAETDETHYFISKELTAAFVAPKESLNT 331
Qy 376 NTTYHCLNEEIKABIEKKY--AKVNSTHYSKYGDLKYFTDGGGLVYVWQPLONRLLDANK 434
Db 332 DPQOTCIKDEFEKIINEVYMSDYNNTYSMNGSYQIFKTTGDLILIQPLVQKSLMFL-- 388
Qy 435 LNNETYSRRROAESITDPMWMTGNGAGGEYSSENSITVAOVAYDNLRIINILE 494
Db 389 ---EQGSEKIRRRDGV-----DVKSRHDIYVQLYQYDILKYDINDALG 431
Qy 495 DLSKAWCREHRAALVNNELSKINPTSVMSMIYNRPVSAKRIGDVISVNSCIVVDQTSVS 554
Db 432 NLAESWCLDQKRTITMLHELKSPSSIVSEVYGRPISAOHLHGDVLAIKSCIEVQSSVQ 491
Qy 555 LHKSRLRLLSA-----SDEKCSRPVPTFKFND--TIYKQGLGVNNEILLTTTLETQEN 609
Db 492 LHKSRRVDAKGVASSETWCYNRPVLTFSFNSTPEVVPGLGLDNEILLGDHRTTEECEIP 551
Qy 610 TEYVFOAKTDMYIYKNYHKLKTVPLSSITTLDTFIALFTLENVDFKVIETLRTDEKRL 669
Db 552 STKIFLGNHAHVYDTHNTSTPIEDIEVLDAFIRUKIDPLENADPKVLDDLSYSPDLSR 611
Qy 670 SNVPDIETMFREYNYAQRVSGLRKDLDDLSTNRNQFVDAFGLSMDLDLGAVGQTVNNAVS 729
Db 612 ANVPDENILREYNSYK---SALYIEAKIATNTPSYNGINSPLQGLGAIGTCLGSGVIS 668
Qy 730 GVATLFSISVTGFNFTNPFPGMLIIVIGVLFYFLTKTKYIETAPIKMIYP--- 786
Db 669 VTAGALGDVGGVVSFLNPFPGGLMLILAIVVVVIIIVFVRQRHSLVSPIDMFPFYAT 728
```

```

Qy 787 -----EIDKLKEREGKSEIAPISEEE 807
Db 729 NPVTTVSSVTGTTVVKTPSVKVDVGGTSA--VSEKE 763

RESULT 37
US-08-720-229-20
: Sequence 20, Application US/08720229
: Patent No. 6022542
: GENERAL INFORMATION:
: APPLICANT: Rose, Timothy M.
: APPLICANT: Bosch, Marnix L.
: APPLICANT: Strand, Kurt
: TITLE OF INVENTION: GLYCOPROTEIN B OF THE RPHV/KSHV
: TITLE OF INVENTION: SUBFAMILY OF HERPES VIRUSES
: NUMBER OF SEQUENCES: 100
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Morrison & Foerster
: STREET: 755 Page Mill Road
: CITY: Palo Alto
: STATE: CA
: COUNTRY: USA
: ZIP: 94304-1018
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/720,229
: FILING DATE: 26-SEP-1996
: CLASSIFICATION: 424
: ATTORNEY/AGENT INFORMATION:
: NAME: Schiff, J. Michael
: REGISTRATION NUMBER: 40,253
: REFERENCE/DOCKET NUMBER: 29938-20002.00
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (415) 813-5600
: TELEFAX: (415) 494-0792
: TELEX: 706141
: INFORMATION FOR SEQ ID NO: 20:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 830 amino acids
: TYPE: amino acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: US-08-720-229-20
```

```

Query Match      23.1%; Score 1050.5; DB 3; Length 830;
Best Local Similarity 32.3%; Pred. No. 3.6e-73;
Matches 244; Conservative 160; Mismatches 293; Indels 59; Gaps 15;

Qy 83 FPRVCSAGVGDFRFQTDHVC--PDASDMVHSEGLLIYKQNIIPMFVRVRYKRVKVT 141
Db 36 YPRFICSIAGTDLMRDRDISCPYKSNAMSGFFIIYKTIETIYTPVRYKKELT 95
Qy 142 STVYNGIYSDSITNQHTFYKSIETPEWETEKMDTIYQCFNSRLN--TGGNLLTYVDRDIN 199
Db 96 QSSYRDVGVVYFLDRVTMGLAMPVYEANLVNSHAQCYSAVAMKRPDGTVFSAFHEDNNKN 155
Qy 200 MTVPLOPVDGVTDPVKRYGSOPELYLEPGFWGYSRRRTTVNCCLMDMPARSPPDPFFV 259
Db 156 NTNLNLPFLNFKSITNKRFTITKEPYFARGPLW--LYSTSTSLNCIVTEATKAKYPSFYA 214
Qy 260 TATGDTVMSPFWGSD--DHENKMKHEKPMFVSVINNYKV--DYON--RGTVPLGKTRIFL 315
Db 215 LTTGEIVGEGPFFNGSKGHFAEPLEK--LTLLENVTMIEDLMNGMNGATTILVRKIAFL 271
Qy 316 DREYTLSEKHLKNNMSYCPDLTLWKAFYNGIQTEHSGSYHFVANDITASTTTSKEDMKEF 375
Db 272 EKADTLFSWEIKEENESVCMKHWTTVTHGLRAETDETHYFISKELTAAFVAPKESLNT 331
```


QY 811 -----IVLAMIHQONSHMETKTRKDPKDSILTRAQNMRLKRSQ--YSNLKNAESVE 860
 Db 823 MIKYSLSVAM---EQOEKKAUKNSGPA-LLASHITNLSLKHGPKYKRLKVNENE 876

RESULT 41

US-08-413-118-2
 ; Sequence 2, Application US/08413118
 ; Patent No. 5688920
 ; GENERAL INFORMATION:
 ; APPLICANT: PAOLETTI, ENZO
 ; APPLICANT: LIMBACH, KEITH J.
 ; TITLE OF INVENTION: NUCLEOTIDE AND AMINO ACID SEQUENCES OF
 ; CANINE HERPESVIRUS 9B, 9C, AND 9D AND USES THEREFOR
 ; NUMBER OF SEQUENCES: 128
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: CURTIS, MORRIS & SAFFORD, P.C.
 ; STREET: 530 FIFTH AVENUE, 25TH FLOOR
 ; CITY: NEW YORK
 ; STATE: NEW YORK
 ; COUNTRY: UNITED STATES OF AMERICA
 ; ZIP: 10036
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent in Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/413,118
 ; FILING DATE: 29-MAR-1995
 ; CLASSIFICATION:
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 08/220,151
 ; FILING DATE: 30-MAR-1994
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: FROMMER, WILLIAM S.
 ; REGISTRATION NUMBER: 25,506
 ; REFERENCE/DOCKET NUMBER: 454310-2670
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (212) 840-3333
 ; TELEFAX: (212) 840-0712
 ; INFORMATION FOR SEQ ID NO: 2:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 879 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: peptide
 ; FRAGMENT TYPE: internal
 ; US-08-413-118-2

Query Match 19.8%; Score 901; DB 1; Length 879;

Best Local Similarity 28.1%; Pred. No. 1.7e-61;
 Matches 252; Conservative 168; Mismatches 402; Indels 76; Gaps 21;

QY 17 LYQALYSLSIAETGVSPPTATWSTESPLTGHGTHDSHGEGNENRDSSEQNKVI 76
 Db 1 MFSLYLIFFIYLIUCOPTPE-STINPLNHLNSTPKPTSDDIRILRESQIE---- 55
 QY 77 YGSPSTFFPYRVGASGVGVDFRQDTHVCPDAS-DMVHSEGIILLYKQNIIPFVRVKY 135
 Db 56 --SDTSTFYMCPPSGSTLVRLPPRACPNYKLGKNTGIAVIFKENISFYKANIY 113
 QY 136 RKVTTSTVYNGIYSDSTINQHTFYKSI-EPWETEKMDTIYQCFNSRLNTGGNLLTYVD 194
 Db 114 YKNIIITTVSGSTVAVITNRYTRDVPICGVPEITELIDRRGCLSKADYIRNNVEFTAFD 173
 QY 195 RDDINMTVLPQVDGVTDPVKYGSQPELYLEPGHFWGSYRRRTVNCELMDMARSNPP 254
 Db 174 KDEEDREVLKPKSKFNTGSGRGWHTVNDTYTKIGGS-GFYHSGTSSVNCIVEVDARSYP 232
 QY 255 FDFVTTATGDTVMSPFPFWSGDD-HENKMKHEKPFVSVINNYKVVDYQNRGTVPILGKTRI 313

Db 233 YDSPAISTGDIHHSPPFGLRGAHTVEISYSTDRFQOIEGYYPIDLDLQLGAPVSRN 292
 QY 314 FLDRREYTLSEKHLKXMSYCPPLWLKAFYNGIQTEHSGSYHFVANDITASF--TTSKED 371
 Db 293 FLTQHTVAVNWPVKIREVCTLAKWREIDRIIRDEYKGSYRFTAKSISATFISDTQFD 352
 QY 372 MKEFNNTVHCLNEBEIKABIEKKY-AKVNSTHSHKYGDLKVFKTGGGLYLVWQPLI----- 424
 Db 353 IDRVKLS-DCAKRAIEAIDKIYKKYKNTHQIQTGELETYLRGGFIIFAPRPMISNEAK 411
 QY 425 -----QNRLLDANKLN-----NETYSRRSRQABSTTDPNMEMTNGAGG 465
 Db 412 LYINELVRSNRTVDLKSLLNPSVRGGARKRSVEENKSKRNIEGGIENVNST----- 465
 QY 466 EYSSENSITVAQVAYDNLRIRNNILEDLSKAWCREQHRRAALVWNLKINPTSVMSM 525
 Db 466 IKTTSVHFAMLOFAYDHIQSHVNEMLSRATATACNLQNKERTLWNEVMKLNPTSVASV 525
 QY 526 IYNRPVSAKRIGDVISVNSCIWVDQTSVSLHKSRLLSASDEKCFSPRPVTFKPMNDSTI 585
 Db 526 AMDQVSARMGLGDVLAVTQCVNISGSSVFQNSMRVL-GSTTTCSYRSLISFKALENSTN 584
 QY 586 Y-KGQLGVNNEILLTITTYLETQCENTEYFOAKTDMYIKNYEHLKTVPLSSITTLDTFI 644
 Db 585 YIEGQGENNELVERKLIETCTANHKRYFKGADYVYFENYAVYRKYPLNEIEMISAYV 644
 QY 645 ALNFTILENVDFKVIELYTRDEKRLSNVFDIETMFREYNYAQRVSGURKOLLDLSTVRN 704
 Db 645 DLNITLDEDFLEVVYTRAELEDTGLLDYSEIQRRNQLHALKFYDI--DSVVKVDNNV 702
 QY 705 QFVDAFGLMDDLGAVGOTVNVNAVSGVATLFSSTVITGFINKPFGGMLLIIVIGLVF 764
 Db 703 VMRGIANFFOGLGDVGAGFGKVLGAANAVIATVSGVSSFLNNPFGALAVGLLLLAGLF 762
 QY 765 AIYFLTKTKIYETAPIKMIYPEIDK-LKE-----REGKSEIAPISSEELER--- 810
 Db 763 AAFLAYRVVSKLKSNNPKALYPVTTKNLKSQVKNNSGNSDGEENDNDIDEKLOQAKE 822
 QY 811 -----IVLAMIHQONSHMETKTRKDPKDSILTRAQNMRLKRSQ--YSNLKNAESVE 860
 Db 823 MIKYSLSVAM---EQOEKKAUKNSGPA-LLASHITNLSLKHGPKYKRLKVNENE 876

RESULT 42

US-08-413-118-3
 ; Sequence 3, Application US/08413118
 ; Patent No. 5688920
 ; GENERAL INFORMATION:
 ; APPLICANT: PAOLETTI, ENZO
 ; APPLICANT: LIMBACH, KEITH J.
 ; TITLE OF INVENTION: NUCLEOTIDE AND AMINO ACID SEQUENCES OF
 ; CANINE HERPESVIRUS 9B, 9C, AND 9D AND USES THEREFOR
 ; NUMBER OF SEQUENCES: 128
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: CURTIS, MORRIS & SAFFORD, P.C.
 ; STREET: 530 FIFTH AVENUE, 25TH FLOOR
 ; CITY: NEW YORK
 ; STATE: NEW YORK
 ; COUNTRY: UNITED STATES OF AMERICA
 ; ZIP: 10036
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent in Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/413,118
 ; FILING DATE: 29-MAR-1995
 ; CLASSIFICATION:
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 08/220,151
 ; FILING DATE: 30-MAR-1994

QY	765	AIYFLTKTKIYETAPIKMIYETDK-LKE	-----REGKSEIAPISEBELER---	811
Db	763	AAFLAYRYVSKLNSPKMAYPVTTKLNKESVKNKNGSNDNDNDEKLOOAKE	822	
QY	811	-----IVLAMHIIHOONSHMETKTRKPKDSILIRAOQNLKFKRS--YSLNKNAESVE	860	
Db	823	MIKYSLSVSAM---EQEHKAIKNSGPA--LLASHITNLSLKHRGPKYKRLKNVNE	876	
<p>RESULT 43</p> <p>US-08-413-118-106</p> <p>Sequence 106, Application US/08413118</p> <p>Patent No. 5688920</p> <p>GENERAL INFORMATION:</p> <p>APPLICANT: PAOLETTI, ENZO</p> <p>APPLICANT: LIMBACH, KEITH J.</p> <p>TITLE OF INVENTION: NUCLEOTIDE AND AMINO ACID SEQUENCES OF</p> <p>TITLE OF INVENTION: CANINE HERPESVIRUS 9B, 9C, AND 9D AND USES THEREFOR</p> <p>NUMBER OF SEQUENCES: 128</p> <p>CORRESPONDENCE ADDRESS:</p> <p>ADDRESSEE: CURTIS, MORRIS & SAFFORD, P.C.</p> <p>STREET: 530 FIFTH AVENUE, 25TH FLOOR</p> <p>CITY: NEW YORK</p> <p>STATE: NEW YORK</p> <p>COUNTRY: UNITED STATES OF AMERICA</p> <p>ZIP: 10036</p> <p>COMPUTER READABLE FORM:</p> <p>MEDIUM TYPE: Floppy disk</p> <p>COMPUTER: IBM PC compatible</p> <p>OPERATING SYSTEM: PC-DOS/MS-DOS</p> <p>SOFTWARE: PatentIn Release #1.0, Version #1.30</p> <p>CURRENT APPLICATION DATA:</p> <p>APPLICATION NUMBER: US/08/413,118</p> <p>FILING DATE: 29-MAR-1995</p> <p>CLASSIFICATION:</p> <p>PRIOR APPLICATION DATA:</p> <p>APPLICATION NUMBER: US 08/220,151</p> <p>FILING DATE: 30-MAR-1994</p> <p>ATTORNEY/AGENT INFORMATION:</p> <p>NAME: FROMMER, WILLIAM S.</p> <p>REGISTRATION NUMBER: 25,506</p> <p>REFERENCE/DOCKET NUMBER: 454310-2670</p> <p>TELECOMMUNICATION INFORMATION:</p> <p>TELEPHONE: (212) 840-3333</p> <p>TELEFAX: (212) 840-0712</p> <p>INFORMATION FOR SEQ ID NO: 106:</p> <p>SEQUENCE CHARACTERISTICS:</p> <p>LENGTH: 879 amino acids</p> <p>TYPE: amino acid</p> <p>STRANDEDNESS: single</p> <p>TOPOLOGY: linear</p> <p>MOLECULE TYPE: peptide</p> <p>FRAGMENT TYPE: internal</p> <p>US-08-413-118-106</p> <p>Query Match 19.8%; Score 901; DB 1; Length 879;</p> <p>Best Local Similarity 28.1%; Pred No. 1.7e-61;</p> <p>Matches 252; Conservative 168; Mismatches 402; Indels 76; Gaps 21;</p>				
QY	17	LYQVALYSLSIAGTGVTPSPNTATWSTESPLTGHVTHDSHGGRGNENRDS	EQKNKI	76
Db	1	MSLYLYIFFIYTLICDPTPE-STINPLNHNLSLTPKPTSDIREILRESQIE	----	55
QY	77	YGSPSTFPYRVCSAGSGDVFRFQTDHVCYCPDAS-DMVHSEGIILLIYKONI	PFMRVRKY	135
Db	56	--SDDTSTFYMCPPSGSLVLEPPRACPNYKLGKNTGEGIAVIFKENSIPYKFKANIY	113	
QY	136	RKVVTSTVYNGIYSDSITNOHTFYKSI-EPWETEKMDTIYQCNSRLNTGGNLLTYVD	194	
Db	114	YKNIITITWVSGSTVAVITNRYTDRVPICVPEITELIDRRGMCLSKADYIRNNYEFTAFD	173	
QY	195	RDDINNTVFLQPDVGTDPVKRYGSOPELYLEPGWFGSYRRRTTVNCELMDMFARSNPP	254	
Db	174	KDDEPREVHLKSKFNTPGSRGWHVNDTYTKIGGS-GFYHSGTSVNCIVEVDARSYP	232	
QY	255	FPFVFTATGDTVEMSPFWSGEDD-HENKMKHEKPFVSVINNYKVVDYQNRGTVP	PLGKTRI	313
Db	233	YDSFAISTGDIHMSPPFGCLRDGAHTEIYSYSTDRFOIEGYYPIDLDTRLQLGAPVSRN	292	
QY	314	FLDRBYTISWEKHLKNMSYCPILTWKAPYNGIQTEHSGSYHFVANDITASF--TTSKED	371	
Db	293	FLTTHQVTVAMNVPKIREVCTLAKWREIDEIIRDEYKGSYRFTAKSISATFISDTTQFD	352	
QY	372	MKEFNTHYCLNEEIKABTEKKY-AKVNSTHSHKYGDLYKFKTDCGLYLVMQPLI-----	424	
Db	353	IDRVKLUS-DCAKREATEADIKYKKYKNTKTHIQTGELETYIARGGFIIFAFPMISNELAK	411	
QY	425	-----QNRLLDAKNKLN-----NETYSRRSRROAESTTDPNEMTGNGAGG	465	
Db	412	LYINELVRNRTVDLKSLLNPSVGRGARKRSVEENKSKNIEGGIENVNNT-----	465	
QY	466	EYSSENSITVAQVQAYDNLRIINILLEDISKAWCREQHRALVWNELSKINFTSVMSM	525	
Db	466	IIKTSSVHFAMQLPAYDHIQSHVNEMLSRATATACNLQNKERTLWNEVMKLNFTSVASV	525	
QY	526	IYRNPVSARKIGDVISVNCIVVDQTSVSLHLSRLLSASDEKCFSPRPVTFKMNDSIT	585	
Db	526	AMDQRVSARMLGDLAVTQCNVIGSSVFIQNSMRVL-GSTTTCYSRPLISFKALENSTN	584	
QY	586	Y-KGOLGVNNEILLTTTLETQENTEYVFOAKTDWYIKNYVHLKTVPLSSITTLDTFI	644	
Db	585	YIEGQLGENNELLVKRLIEPTCAHAKRYFKGADYVYFENYAVYKRVPLNEIEMISAYV	644	
QY	645	ALNFTLLENVDFKVELYTRDEKRLSNVPDIETMFREYNYAQRVSGLRKDLLDLSTNRN	704	
Db	645	DLNITLLEDREFLEPVYTRAELEDGLLDYSEIQRRNLHAKFYDI--DSVVKVDNVV	702	
QY	705	QFVDAFGSLMDLGAQGVTVNNAVSGVATLFESSIVTGFINIKNPPFGMLMIIVIGVLF	764	
Db	703	VIMRGIANTFQGLDGVAGFGKVLGAANAVIATVSGVSSFLNPNPFGALVGLLILAGLF	762	

```

Db 174 KDEDREVLHPSKFNTPGSRGWHTVNDYTKIGGS-GFYHSGTSVNCIVEEDARSVYP 232
QY 255 PFFFVTATGDTVMSPFWSGEDD-HENKMHKPMFVSVNNYKVVDYQNRGTVPLGKTRI 313
Db 233 YDSFALSTGDIITHMPPFGLRGATEYISYSTDFOQIEGYPIDLDTLRLQAGAPVSRN 292
QY 314 FLDRREYTLSEKHLKNNMSYCPPLTLWKAFCYNGIQTEHSGSYHFVANDITASF--TTSKED 371
Db 293 FLTQHTVTVMNWVPKIREVCTLAKWREIDEIIRDEYKGSYRFTAKSISATISDTTQFD 352
QY 372 MKEFNNTYHCLNEBEIKAEIKKY-AKVNSTHSGYDLYKFKTDGGLYLWQPLI----- 424
Db 353 IDRVKLS-DCAKREAIEADKIYKKYKNTKTHIQTELETYLARGGFFIAFRPMISNELAK 411
QY 425 -----QNRLLDAKNKLN-----NETYSRRSRQAESTTDPMMEMTNGAGG 465
Db 412 LYINELVSRNRTVDLKSLLNPSVRGARGKRVSVEENKSRKNIEGGIENVNST----- 465
QY 466 EYSSENSITVAQVAYDNLRIINNILEDLSKAWCREQHRAALVWNLKINPTSVMSM 525
Db 466 IIKTTSSVHFAMLOFAYDHIQSHVNEMLSRATAMCNLQNKERTLWNEVMKLNPTSVASV 525
QY 526 IYNRPVSAKRIGDVISVNCIVDQTSVLSLHLSASDEKCFSPRPVTFKFNWDSSTI 585
Db 526 AMDQVSARMGLDGLAVTQCVNISGSSVFIQNSMRVL-GSTTTCYSRPLISFKALENSTN 584
QY 586 Y-KGOLGVNNEILLTTTLETQCENTEYVFOAKTDMYIKNYEHLKTVPLASITLDTFI 644
Db 585 YIEGOLGENNELVERKLEIPTCTANHRYFKFGADYVVFENYAYVKRPLNEIEMISAYV 644
QY 645 ALNFTLLENVDFKVIELYTRDEKRLSNVFDIETMFREYNYAQRVSGLRKOLLDLSTNRN 704
Db 645 DLNITLLEDRFLPLEVYTRAELEDGLDYSEIQRRNLHALKFYDI--DSVVKVDNNV 702
QY 705 QFVDAFGLSMDLGLAVGTQVNVAVSGVATLFSSIVTGFINFKNPFGLMGLIIVIGVLF 764
Db 703 VIMRGIANFFQGLGVGAGFGKVLGAANAVIATVSGVSSFLNPPFGALAVGLLILAGLF 762
QY 765 AIYFLTKTKIYETAPIKMIYPIEDK-LKE-----REGKSEIAPISEEELER--- 810
Db 763 AAFLAYRVSVKLSNPMKALYPVTTKNLKESVNGNSGNSDGEENDNDIIEELQOQAKE 822
QY 811 -----IVLAMHIHQNSHMETKTRKDPKDSILTTRAQNMRLKRSQ--YSNLKNABSVE 860
Db 823 MIKWSLVSAM---EQQEHKAIKKNSGPA-LLASHITNLSLKHRCGPYKRLKNNVENE 876

RESULT 44
US-08-473-446-2
; Sequence 2, Application US/08473446
; Patent No. 6017542
; GENERAL INFORMATION:
; APPLICANT: PAOLETTI, ENZO
; APPLICANT: LIMBACH, KEITH J.
; TITLE OF INVENTION: NUCLEOTIDE AND AMINO ACID SEQUENCES OF
; TITLE OF INVENTION: CANINE HERPESVIRUS gB, gC, AND gD AND USES THEREFOR
; NUMBER OF SEQUENCES: 128
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CURTIS, MORRIS & SAFFORD, P.C.
; STREET: 530 FIFTH AVENUE, 25TH FLOOR
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: UNITED STATES OF AMERICA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/473,446
; FILING DATE:
; CLASSIFICATION:

```

```

; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/413,118
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: FROMMER, WILLIAM S.
; REGISTRATION NUMBER: 25,506
; REFERENCE/DOCKET NUMBER: 454310-2670
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 840-3333
; TELEFAX: (212) 840-0712
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 879 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FRAGMENT TYPE: internal
; US-08-473-446-2

```

Query Match 19.8%; Score 901; DB 3; Length 879;

Best Local Similarity 28.1%; Pred. No. 1.7e-61;

Matches 252; Conservative 168; Mismatches 402; Indels 76; Gaps 21;

```

QY 17 LYQVALYSLSLAETGVTSPNPTATWSTESPLTGHGTHDSHGGRGNENNRDSEBQNKNI 76
Db 1 MFSLYLIFFIYLIICDPTTPE-STINPLNHHNLSTPKPTSDDIREILRESQIE----- 55
QY 77 YGSPSTPYRVCSAGVDVRFOTDHWCPDAS-DMVHSEGIILYKNIIPFVRVKRY 135
Db 56 --SDTSTFYMCPPSPGSTLVRLEPPRACPNYKLGKNFTGEGIAVIFKENISPYKFKANIY 113
QY 136 RKVVTTSVNYGIVSDSTNQHTFYKSI-EPWETEKMDTIYQCFNSLRNLGGLLTYYVD 194
Db 114 YKNIITTVSGSGTYAVITNRYTRDVRPIGVPEITELIDRRGMCLSKADYIRNNVEFTAFD 173
QY 195 RDDINMTVFQPDVGVTDPVKRYGQPELYLEPGWFGWSYRRRTTVNCELMDMFARSNPP 254
Db 174 KDEDREVLHPSKFNTPGSRGWHTVNDYTKIGGS-GFYHSGTSVNCIVEEDARSVYP 232
QY 255 PFFFVTATGDTVMSPFWSGEDD-HENKMHKPMFVSVNNYKVVDYQNRGTVPLGKTRI 313
Db 233 YDSFALSTGDIITHMPPFGLRGATEYISYSTDFOQIEGYPIDLDTLRLQAGAPVSRN 292
QY 314 FLDRREYTLSEKHLKNNMSYCPPLTLWKAFCYNGIQTEHSGSYHFVANDITASF--TTSKED 371
Db 293 FLTQHTVTVMNWVPKIREVCTLAKWREIDEIIRDEYKGSYRFTAKSISATISDTTQFD 352
QY 372 MKEFNNTYHCLNEBEIKAEIKKY-AKVNSTHSGYDLYKFKTDGGLYLWQPLI----- 424
Db 353 IDRVKLS-DCAKREAIEADKIYKKYKNTKTHIQTELETYLARGGFFIAFRPMISNELAK 411
QY 425 -----QNRLLDAKNKLN-----NETYSRRSRQAESTTDPMMEMTNGAGG 465
Db 412 LYINELVSRNRTVDLKSLLNPSVRGARGKRVSVEENKSRKNIEGGIENVNST----- 465
QY 466 EYSSENSITVAQVAYDNLRIINNILEDLSKAWCREQHRAALVWNLKINPTSVMSM 525
Db 466 IIKTTSSVHFAMLOFAYDHIQSHVNEMLSRATAMCNLQNKERTLWNEVMKLNPTSVASV 525
QY 526 IYNRPVSAKRIGDVISVNCIVDQTSVLSLHLSASDEKCFSPRPVTFKFNWDSSTI 585
Db 526 AMDQVSARMGLDGLAVTQCVNISGSSVFIQNSMRVL-GSTTTCYSRPLISFKALENSTN 584
QY 586 Y-KGOLGVNNEILLTTTLETQCENTEYVFOAKTDMYIKNYEHLKTVPLASITLDTFI 644
Db 585 YIEGOLGENNELVERKLEIPTCTANHRYFKFGADYVVFENYAYVKRPLNEIEMISAYV 644
QY 645 ALNFTLLENVDFKVIELYTRDEKRLSNVFDIETMFREYNYAQRVSGLRKOLLDLSTNRN 704
Db 645 DLNITLLEDRFLPLEVYTRAELEDGLDYSEIQRRNLHALKFYDI--DSVVKVDNNV 702
QY 705 QFVDAFGLSMDLGLAVGTQVNVAVSGVATLFSSIVTGFINFKNPFGLMGLIIVIGVLF 764

```

Db 703 VIMRGIANFFQGLGDVGAGFKVVLGAANAVIATVSGVSSFLNPFALAVGLLILAGLF 762
Qy 765 AIYFLTKTKIYETAPIKMIYPEIDK-LKE-----REGKSEIAPISEBELER--- 810
Db 763 AAFLAYRYVSKLSPKALYPVTTKNLKEVKNNGSNGNSDGEENDNDIEBKLOAKE 822
Qy 811 -----IVLAMHIHQONSHMETKTRKDPKDSILTRAQNMRLKRSQ--YSNLKNAESVE 860
Db 823 MIKYSLSVSAM---EQOEHAIKKNSGPA-LLASHITNLSLKHGPKYKRLKNVNE 876

RESULT 45
US-08-473-446-3
; Sequence 3, Application US/08473446
; Patent No. 6017542
; GENERAL INFORMATION:
; APPLICANT: PAOLETTI, ENZO
; APPLICANT: LIMBACH, KEITH J.
; TITLE OF INVENTION: NUCLEOTIDE AND AMINO ACID SEQUENCES OF
; TITLE OF INVENTION: CANINE HERPESVIRUS 9B, 9C, AND 9D AND USES THEREFOR
; NUMBER OF SEQUENCES: 128
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: CURTIS, MORRIS & SAFFORD, P.C.
; STREET: 530 FIFTH AVENUE, 25TH FLOOR
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: UNITED STATES OF AMERICA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/473,446
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/413,118
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: FROMMER, WILLIAM S.
; REGISTRATION NUMBER: 25,506
; REFERENCE/DOCKET NUMBER: 454310-2670
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 840-3333
; TELEFAX: (212) 840-0712
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 879 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FRAGMENT TYPE: N-terminal
US-08-473-446-3

Query Match 19.8%; Score 901; DB 3; Length 879;
Best Local Similarity 28.1%; Pred. No. 1.7e-61;
Matches 252; Conservative 168; Mismatches 402; Indels 76; Gaps 21;
Qy 17 LYQVALYSLSIAETGVTSPPNTATWSTESPLTHGYGTHDSHGSGNGNENRDSBEQKN 76
Db 1 MFLSLYLIFFIYLIICDPTPE-STINPLNHLNLSLTPKTSDDIRILRESQIE--- 55
Qy 77 YGSPSTPEPYVCSAGVGDFVRFOTDHCVPDAS-DMVHSGILLIYKONIPFMRVKY 135
Db 56 --SDDSTFWCPCPPSGSTLVRLPPRACPNYKLGKNFTGEGIAVFKENISPYKFKANIY 113
Qy 136 RKVVTTSVNGIYSDITNQHTFYKSI-BPWETEKMDTIYQCFSNLSRLNTGGNLLTYVD 194
Db 114 YQNIIITVWSGSGTYAVITNRYTRDVPVIGVPEITELIDRGMCLSKADYIRNNVEFTAFD 173

Qy 195 RODINMTVLOPVDGVTDPVKRYGSOPELYLEPGFWGYSRRRTTYNCMLMDMFARSNPP 254
Db 174 KOEDPREVHLKSKFNTPGSRGWHITVNDYTKIGGS-GFYHSOTSVNCIVEEDAKSVYP 232
Qy 255 FDFFTVATGDTVEMSPFWSGEDD-HENKHEKPMFWFVSVINNYKVVDYQNRGTVPGLGTRI 313
Db 233 YDSFALSTGDIIMHSPFGLRDCAHTEYISYSTDRFQQIEGYYPIDLDLTLQLGAPVRN 292
Qy 314 FLDREYTTLSWEKHLKNMSYCPITLWKAFYNGIQTEHSGSYHFAVNDITASF--TTSKED 371
Db 293 FLTTQHTVTAVNWPVKIREVCTLAKWREIDEIIRDEYKGSYRFTAKSISATFISDTQPD 352
Qy 372 MKEFNTTYHCLNEEIKAEIEKKY-AKVNSTHSGYGLKYFKTKDGLGLYLVNQPLI----- 424
Db 353 IDRVKLS-DCAKREATEADKIYKKYKNTHTQTGLEYTLARGGFIIFAPRPMISLAK. 411
Qy 425 -----QNRLLDAKNKLN-----NETYSRRRRQAESTTDPMMMTGNGAGG 465
Db 412 LYINELVRSNRITVDLKSLLNPSVRGARGKRSVEENKRSKRNIEGGIENVNST----- 465
Qy 466 EYSSENSITVAQVQYAYDMLRIRINNIBDLKAWCREQHRALVWNELSKINPTSVMSM 525
Db 466 IIKTSSVHFAMLQFAYDHIQSHVNEMLSRIATAWCNQLNKERTLNWNEVMKLNPTSVASV 525
Qy 526 IYNRPVSAKRIGDVISVNCIVVDQTSVSLHLSKLLSASDEKCFRPPVPTFKFMDSTI 585
Db 526 AMDQVRSARMLGDVLAVTQCVNISGSSVFIONSMRVL-GSTTTCYSRPLISFKALENSTN 584
Qy 586 Y-KGOLGVNNEILLTTTILETCOENTEYFYFOAKTDMYIYKNYEHKLTVPPLSSITTLDTFI 644
Db 585 YIEGQLGNNELLVERKLEPCTANHRYKFKGADYVYFENYAVYKVPLENEIMISAYV 644
Qy 645 ALNFTLLENVDFKVIELYTRDEKRLSNVFDIETMPREYNYAQRVSGLRKDLDLSTNRN 704
Db 645 DLNITLLEDFLEPLLEVYTRAELEDGLDYSEIQRNQLHALKPYDI--DSVVKVDNNV 702
Qy 705 QFVDAFGSLMDDLGAVGQTVNNAVSGVATLPSISIVTGFINFKNPFGGMLMIIVIGVLF 764
Db 703 VIMRGIANFFQGLGDVGAGFKVVLGAANAVIATVSGVSSFLNPFALAVGLLILAGLF 762
Qy 765 AIYFLTKTKIYETAPIKMIYPEIDK-LKE-----REGKSEIAPISEBELER--- 810
Db 763 AAFLAYRYVSKLSPKALYPVTTKNLKEVKNNGSNGNSDGEENDNDIEBKLOAKE 822
Qy 811 -----IVLAMHIHQONSHMETKTRKDPKDSILTRAQNMRLKRSQ--YSNLKNAESVE 860
Db 823 MIKYSLSVSAM---EQOEHAIKKNSGPA-LLASHITNLSLKHGPKYKRLKNVNE 876

Search completed: October 28, 2003, 15:30:06
Job time : 42 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 28, 2003, 15:27:11 ; Search time 86 Seconds
(without alignments)
1596.494 Million cell updates/sec

Title: US-10-055-364-24
Perfect score: 4547
Sequence: 1 MAGSLKRLGSLVLAFLYQV.....KRGYSNLKNAESVEMLNTL 865

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_19Jun03.*

1: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*
24: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	4547	100.0	865	22 AAB31679	Amino acid sequence
2	1937	42.6	845	18 AAW26475	KSHV glycoprotein
3	1927.5	42.4	829	21 AAB53128	Macaca mulatta rha
4	1840	40.5	856	22 AAU14094	Peptide sequence f
5	1798.5	39.6	927	8 AAP70768	Egstein-Barr virus
6	1128	24.8	907	15 AAP63462	Human cytomegalovi
7	1123	24.7	907	18 AAW27274	Human cytomegalovi
8	1121.5	24.7	906	8 AAP70289	Sequence encoded b
9	1114	24.5	906	10 AAP91047	gB envelope protei

10	1002	22.0	195	22	AAB31678	A viral glycoprote
11	901	19.8	879	16	AAR77024	Canine herpesvirus
12	876.5	19.3	865	14	AAR30169	Marek's Disease Vi
13	851.5	18.7	933	13	AAR27807	Bovine herpes viru
14	851.5	18.7	933	13	AAR41343	Bovine herpesvirus
15	847.5	18.6	904	12	AAR41465	HSVgB polypeptide.
16	845.5	18.6	903	8	AAP71135	Herpes Simplex Vir
17	844	18.6	904	23	AAB17812	Herpes simplex vir
18	843.5	18.6	903	8	AAP70426	Recombinant herpes
19	841	18.5	904	17	AAW00375	HSV-1 glycoprotein
20	841	18.5	904	22	AAB74441	Herpes simplex vir
21	839	18.5	854	19	AAW72113	HSV-2 strain SB5 C
22	839	18.5	904	14	AAR41778	Glycoprotein B (gB
23	839	18.5	904	17	AAW00376	HSV-2 glycoprotein
24	839	18.5	904	19	AAW72193	HSV-2 strain SB5 C
25	839	18.5	904	22	AAB74442	Herpes simplex vir
26	838.5	18.4	903	7	AAP60244	Herpes simplex vir
27	838	18.4	885	17	AAR92747	SAB virus gB glyco
28	838	18.4	904	18	AAW34552	Herpes simplex vir
29	838	18.4	907	8	AAP71136	Herpes Simplex Vir
30	836.5	18.4	904	14	AAR41779	Glycoprotein B (gB
31	836	18.4	795	19	AAW72062	HSV-2 strain SB5 C
32	833	18.3	904	9	AAP80914	Sequence of Herpes
33	832	18.3	868	13	AAR22615	Varicella-zoster v
34	829.5	18.2	904	18	AAW34553	Herpes simplex vir
35	823.5	18.1	905	9	AAP80915	Sequence of Herpes
36	821	18.1	873	16	AAR79459	Infectious Laryngo
37	819.5	18.0	973	8	AAP70769	Glycoprotein B of
38	818.5	18.0	891	17	AAR92746	B virus gB glycopr
39	818.5	18.0	891	19	AAW70293	Simian herpesvirus
40	811	17.8	854	8	AAP70347	Varicella-zoster v
41	808.5	17.8	903	6	AAP50312	Herpes simplex vir
42	789.5	17.4	943	21	AAV32470	DNA encoding felin
43	789.5	17.4	943	22	AAW50113	Feline herpesvirus
44	777.5	17.1	913	19	AAW68404	Aujeszky's disease
45	757	16.6	1086	18	AAW36051	Hybrid Marek's dis

ALIGNMENTS

RESULT 1

AAB31679
ID AAB31679 standard; Protein; 865 AA.

AC AAB31679;

DT 30-APR-2001 (first entry)

DE Amino acid sequence of a viral glycoprotein B envelope protein.

DE Glycoprotein B; envelope protein; vaccine; passive immunity; pig.

OS Swine gamma herpesvirus.

PN WO200103733-A1.

PD 18-JAN-2001.

XX 07-JUL-2000; 2000WO-US18647.

XX 08-JUL-1999; 99US-0142736.

PR 02-DEC-1999; 99US-0168532.

XX (BIOT-) BIOTRANSPLANT INC.

XX Patience C;

XX MPI; 2001-103033/11.

XX N-PSDB; AAF25199.

XX New polynucleotide encoding a polypeptide corresponding to porcine
PT gamma-herpesvirus glycoprotein B (pGHV-gpB) for immunizing a pig

PT against a porcine gamma-herpesvirus -

PS Claim 9; Fig 3; 82pp; English.

XX The present sequence represents a swine gamma herpesvirus glycoprotein
CC B envelope protein. Probes and primers derived from DNA encoding the
CC glycoprotein B envelope protein are used for detecting the presence
CC of swine gamma herpesvirus. Antibodies directed against the envelope
CC protein are also useful for detecting presence of the virus. The
CC envelope protein can be used as a vaccine for creating passive
CC immunity in a pig.

XX Sequence 865 AA;

Query Match 100.0%; Score 4547; DB 22; Length 865;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 865; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAGSLKLRGSLVLAALWLYQVALYSLSTAETGVTSPPTATWSTESPLTGHYTHDSSHGE 60
DB 1 MAGSLKLRGSLVLAALWLYQVALYSLSTAETGVTSPPTATWSTESPLTGHYTHDSSHGE 60
QY 61 RGNENRDSBQNKNIYGSPTFPYRVCSAGSGVDVFRFQTDHVCPSDASDMVHSEGI 120
DB 61 RGNENRDSBQNKNIYGSPTFPYRVCSAGSGVDVFRFQTDHVCPSDASDMVHSEGI 120
QY 121 YQNIIIPFMRVRKRVKVTSTVYNGIYSDSITNQHTFYKSIPEWETEKMDTIYQCFNS 180
DB 121 YQNIIIPFMRVRKRVKVTSTVYNGIYSDSITNQHTFYKSIPEWETEKMDTIYQCFNS 180
QY 181 LRLNTGNLTYYVDRDINDMTVFLQPDVGTTPDKRYGSOPELYLEPGWFGSVRRRTTV 240
DB 181 LRLNTGNLTYYVDRDINDMTVFLQPDVGTTPDKRYGSOPELYLEPGWFGSVRRRTTV 240
QY 241 NCELMDFARSNPPDFEFTVATGTVEVMSFPWGSDDHKNMKHEKPFVSVINNKVVDY 300
DB 241 NCELMDFARSNPPDFEFTVATGTVEVMSFPWGSDDHKNMKHEKPFVSVINNKVVDY 300
QY 301 QNRGTVPGLKTRIFLDREEVYLSWEKHLKMSYCPFLWKAIFYNGIQTESGSHYFVAND 360
DB 301 QNRGTVPGLKTRIFLDREEVYLSWEKHLKMSYCPFLWKAIFYNGIQTESGSHYFVAND 360
QY 361 ITASFTTSKEDMKFNTTYHCLNEIKAEIEKKYAKVNSTHSHYKGLYKPTDGLGLVW 420
DB 361 ITASFTTSKEDMKFNTTYHCLNEIKAEIEKKYAKVNSTHSHYKGLYKPTDGLGLVW 420
QY 421 QPLIQNRLDAAKNNLTYSRRRQAEISTDPMEMTNGAGGEYSSENSITVAQVY 480
DB 421 QPLIQNRLDAAKNNLTYSRRRQAEISTDPMEMTNGAGGEYSSENSITVAQVY 480
QY 481 AYDNLRIINNILEDLSKACREOHRALVWNLKINPTSVMSMIYNRVPSAKRIGDVI 540
DB 481 AYDNLRIINNILEDLSKACREOHRALVWNLKINPTSVMSMIYNRVPSAKRIGDVI 540
QY 541 SVSNICVVDQTSVSLHSLRLSASDEKCFRPPVTPKFWNDSTIYKQGLGVNNEILLTT 600
DB 541 SVSNICVVDQTSVSLHSLRLSASDEKCFRPPVTPKFWNDSTIYKQGLGVNNEILLTT 600
QY 601 TYLETCENTEEYFQAKTDMYIKNYEHLKTVPLSSITTLDTFTALNFTLLENVDFKVE 660
DB 601 TYLETCENTEEYFQAKTDMYIKNYEHLKTVPLSSITTLDTFTALNFTLLENVDFKVE 660
QY 661 LYTRDEXRLSNVFIETMFREYNYAQRVSGRLKDLDLSTNRNQFVDAFGSLMDDLGA 720
DB 661 LYTRDEXRLSNVFIETMFREYNYAQRVSGRLKDLDLSTNRNQFVDAFGSLMDDLGA 720
QY 721 GQTVNAVSGVATLFFSSIVTGFNFINKPFGGMLIIVLGVLPAIYFLTKTKIYETAP 780
DB 721 GQTVNAVSGVATLFFSSIVTGFNFINKPFGGMLIIVLGVLPAIYFLTKTKIYETAP 780
QY 781 IKMIYPIDKLKREGKSEIAPISEEELERIVLAMHIHQONSHMETKTRDKPDKSILTRA 840
DB 781 IKMIYPIDKLKREGKSEIAPISEEELERIVLAMHIHQONSHMETKTRDKPDKSILTRA 840

QY 841 QNMLRKRSYGYSNLKNAESVEMLNTL 865
DB 841 QNMLRKRSYGYSNLKNAESVEMLNTL 865

RESULT 2

AW26475
ID AAW26475 standard; Protein; 845 AA.

XX AC AAW26475;

XX 02-JAN-1998 (first entry)

XX KSHV glycoprotein B.

XX KSFV; gamma herpes virus; glycoprotein B; vaccine; infection;
KW diagnosis.

XX Human Kaposi's sarcoma-associated herpesvirus.

XX Key Location/Qualifiers

FT Misc-difference 841
FT /label= Leu, Pro
FT /note= "encoded by CYG"

XX WO9712042-A2.

PN

XX 03-APR-1997.

XX 26-SEP-1996; 96WO-US15702.

XX 26-SEP-1995; 95US-0004297.

XX (UNIW) UNIV WASHINGTON.

XX Bosch ML, Rose TM, Strand K;

XX WPI; 1997-212901/19.

XX N-PSDB; AAT84648.

XX DNA encoding glycoprotein B of retroperitoneal fibromatosis and

XX Kaposi's sarcoma associated herpes viruses - useful in vaccines for

XX treatment of herpes infection or for detection of viral DNA

XX Claim 11; Fig 19; 138pp; English.

XX This polypeptide comprises the full-length glycoprotein B (GB) of

XX human Kaposi's sarcoma-associated herpes virus (KSHV). It is the

XX encoded product of a consensus sequence (see AAT84648) obtained from

XX Kaposi's sarcoma tissue DNA. Herpes viruses GB molecules are

XX associated with fibroproliferative and neoplastic conditions in

XX primates, including humans. Claimed GB polypeptides (AAW26473-76)

XX can be expressed in transformed host cells for use in vaccines, for

XX the treatment of herpes virus infection, to assess herpes virus

XX infection, to design or screen antiviral agents and to raise

XX antibodies. Claimed antigenic peptides (see AAW26477-86) derived

XX from KSHV and/or retroperitoneal fibromatosis-associated herpes virus

XX can be used to raise monoclonal or polyclonal antibodies.

XX Sequence 845 AA;

Query Match 42.8%; Score 1937; DB 18; Length 845;
Best Local Similarity 48.5%; Pred. No. 4.8e-134;
Matches 382; Conservative 135; Mismatches 237; Indels 34; Gaps 10;

QY 83 FPYRVCSAGSGVDVFRFQTDHVCPSDASDMVHSGILLIYKONTIIPFMRVRKRVKVTTS 142

DB 63 YQFRVCSAITGELFFNLNLEQTCPTDKYHQBGLLVYKKNLIVPHIFKRVRRKRTATS 122

QY 143 TVYNGIYSDSITNQHTFYKSIPEWETEKMDTIYQCFNSLRNTGGNLLTYVDRDDINMTV 202

DB 123 TVYRGLTESAITNKYELPRVPVLYEISHMDSTYQCFSSMKVNVNGENTFTDSDVNTTV 182

```

QY 203 FLQPDGVTPDVKRYGSGQPELYLEPGWFMGYSYRRRTTNCVCELMDFARSNPPDFPVAT 262
DB 183 FLQPDVGLTNDIQRYFSQPIVAEPGWPFGIYVRVTNCEIVDMARSAPYNYFVTS 242
QY 263 GDTVEMSPWSEDDHKNHKEKWP - VSVINNYKVVDYQNRGTVPGLKTRIFLDREBYT 321
DB 243 GDTVEVSPFCYNSSCSCTPSSNKNGLSVQVLNHTVTVSDRGTSPTQNRIFVETGAYT 302
QY 322 LSWKHKLKMSYCPITLWKAIFYNGIQTEHSGVGHFVANDITASFTTSKEDMKEFTVHC 381
DB 303 LSWASEKTTAVCPALMKTFRPSIQTHEDSFHFVANEITATFAPTPVANFTDTYSC 362
QY 382 LNEEIKAEIEKKYAKVNSHYSYKGLYFKTDGGLYVWQPLIQRLDADKNNKNET-- 439
DB 363 LSTDINTTLNASKAKLASTHVPNGTVQYFHTTGLYLVWQPSAINLTHAQDSCGNPTSS 422
QY 440 -----YSRRSRQAESTTDPMMEMTONGAGGEYSSENSITVAQVQVAYNLIR 488
DB 423 PPSASPMTTSASRRKRSASTA-----AAGGGGSTDN-LSYTLQAFAYDKLRDG 471
QY 489 INNILEDLSKAWCROHRAALVWNLKSKINPTSVMSIYNRPVSASAKRIGDVISVNCIIV 548
DB 472 INQVLEELSRACROVRDNLWYELSKINPTSVMTAIYGRVSAKFGVDALISVTECINV 531
QY 549 DQTSVSLHKLRLLSASDEKCFSPRPVTFKFMNDSTIYKQGLGVNNEILLTLYLETCOE 608
DB 532 DQSSVNIHKSAR--TNSKDVCYARPLVTFKFLNSSNLFTGOLGARNEILLTNNOVETCKD 589
QY 609 NTEYFQAKTDYIYKVNHEHLKTVPLSSITLDTFIALNFTLLENVDPKVIELYTRDEKR 668
DB 590 TCEHYFTRNEFLVKDYAVLRTINTDITSLNFTALNLSFIQNDKAIELYSABER 649
QY 669 L-SNVFDETMEPREVNYAQRVSGLRKOLL-D-LSNRNQFVDAPGSLMDDLGAVGQTVN 726
DB 650 LASSVFDLETMEPREVNYTHRLAGURELDNTDMNKEFERVLDSEIIVADLGGIGKTVN 709
QY 727 AVSGVATLFPSSIVTGFINFKPPGGLMIIVIGVLFALYFLTKTKIYETAPIKMIYP 786
DB 710 VASSVVTLCGLSVTGFINFKHPLGLGMLIIVIAILLFMLSRTNTIAQAPVKMIYP 769
QY 787 EIDKLKEREKGEIAPISEEELERVLAMHHQNSHMETKTRDKPSILTRAQNMURK 846
DB 770 DVDR---RAPPSGGAP-TREEIKNILLGWHQLQOEEROKADDLKKSTPVSFORTANGLRQ 825
QY 847 R-SGYSNL 853
DB 826 LURGYKPL 833

RESULT 3
AAB53128
ID AAB53128 standard; Protein; 829 AA.
XX
AC AAB53128;
XX
DT 28-FEB-2001 (first entry)
XX
DE Macaca mulatta rhadinovirus 17577 RRV ORF8 protein SEQ ID NO:13.
XX
KW Macaca mulatta rhadinovirus 17577; RRV; rhesus macaque rhadinovirus;
KW genome; Kaposi's sarcoma-associated herpesvirus; KSHV; interleukin 6;
KW IL-6; macrophage inflammatory protein; MIP; diagnosis; vaccine;
KW cytostatic; anti-HIV; gene therapy; infection; Kaposi's sarcoma;
KW lymphoproliferative disorder; B-cell hyperplasia; lymphadenopathy;
KW splenomegaly; hypergammaglobulinemia; autoimmune haemolytic anaemia.
XX
OS Macaca mulatta rhadinovirus 17577.
XX
PN WO200028040-A2.
XX
PD 18-MAY-2000.
XX

```

```

PF 05-NOV-1999; 99WO-US26260.
XX
PR 06-NOV-1998; 98US-0107507.
PR 20-NOV-1998; 98US-0109409.
XX
XX (UYOR-) UNIV OREGON HEALTH SCI.
XX
PI Wong SW, Axthelm MK, Searles RP;
XX
XX WPI; 2000-376552/32.
XX
PT New rhesus rhadino virus for producing non-human primate model useful
PT for testing potential treatments and efficacy of the candidate vaccine
PT for conditions associated with RRV infection .
XX
PS Claim 5; Page 123-124; 141pp; English.
XX
CC The present invention describes a novel rhesus macaque rhadinovirus
CC called macaca mulatta rhadinovirus 17577 (RRV). AAC64754 represents the
CC RRV genome sequence, and AAB53123 to AAB53204 represent the proteins
CC encoded by the genome sequence. The present invention also specifically
CC claims the individual open reading frame (ORF) nucleotide sequences from
CC the genome which encode the individual proteins, but these sequences are
CC not given. A non-human animal infected with RRV can be used for testing
CC the efficacy of drug in the treatment of condition associated with
CC infection with RRV such as Kaposi's sarcoma, lymphoproliferative
CC disorders, B-cell hyperplasia, lymphadenopathy, splenomegaly,
CC hypergammaglobulinemia or autoimmune haemolytic anaemia, by
CC administering the drug to an immuno-compromised non-human primate
CC preferably Rhesus macaque monkey obtained by as a result of infection
CC by Simian Immunodeficiency Virus (SIV). RRV is useful for producing
CC non-human primate model for testing potential treatments for conditions
CC associated with RRV infection. It is also useful for testing the
CC efficacy of the candidate vaccine against RRV infection or conditions
CC associated with its infection by administering the vaccine to the
CC subject capable of infection with RRV, inoculating the subject with RRV
CC and observing the effect of vaccine. AAC64755 to AAC64765 and AAB53205
CC to AAB53213 represent sequence used in the exemplification of the
CC present invention.
XX
SQ Sequence 829 AA;
XX
Query Match 42.4%; Score 1927.5; DB 21; Length 829;
Best Local Similarity 44.4%; Pred. No. 2.3e-133;
Matches 388; Conservative 146; Mismatches 258; Indels 81; Gaps 14;
QY 11 VLALWLYQVALYSLSIATGVTSPPTATWSTESPLTGHYTHDSSHGGRGNENRDSE 70
DB 10 LLRAWVV--IIAIGTAVGE-NVITPKGATTAKTP----- 42
QY 71 EQKNVIYGSPT-----FPYVCSASGVDFRFRQTDVHVCPSADSMVHSEGIL 118
DB 43 -----GPSTPTPPENPRAEAFKFRVCSASATGELFRNLEKTCPTGTDKTHQEGIL 94
QY 119 LIYKQNIIPFMRVRKRVKVVTTSTVYNGIYSDSTNQHTFYKSIPEWTEKMDTIYOCF 178
DB 95 MVFKKNIIVPIFKVRVRYRKVATSVTVRGVTETAVTGKQEVIRVPVQYENHMDITTYOCF 154
QY 179 NSLRANTGGNLLTYVDRDDINMTVFLQPDGVTPDVKRYGSGQPELYLEPGWFMGYSYRRRT 238
DB 155 SSMRVNVNGIIVNTYDRTDFTNQTVFLQPDVGLTNDIQRYFSQPIVYTPGWFPGIYVRVT 214
QY 239 TVNCELMDMFARSNPPDFPVATGDTVEMSPWSE-----DDHKNHKEKWPVSVI 292
DB 215 TVNCEIVDMTARSAPYSYFVTALGDTVEVSPFCYNSSCSCTPSSNKNGLSVQVLNHTVTVSDRGTSPTQNRIFVETGAYT 269
QY 293 NNYKVDYQNRGTVPGLKTRIFLDREBYTLSWEKHLKMSYCPITLWKAIFYNGIQTEHSG 352
DB 270 TNYTIVDFATR--QPTTETRVFADSGEYTVSWKAEDPKSAVCALTWKTFPRAIQTHTHEA 327
QY 353 SYHFVANDITASFTTSKEDMKEFTNTYHCLNEEIKAEIEKKYAKVNSHYSYKGLYFKT 412
DB 328 SYHFVANDVTATFTSPLSQVNTNFTGTYPCLNDVDTQTLNATIKLSDSHTATNGSEQYVET 387

```

QY 413 DGGILVWQPLQIRLLDAKNNLNNETYSSRRRQAESTTDPMMEMTNGAGGEYSSENS 472
 Db 388 EGGFLWQPLTPLSLADEWRELNGTT-----PAPPTTTANRVRRVSGVNEQATDD 440
 QY 473 ITVAQVQAYDNLIRINNILEDLKAWCREQRAALVWNLKSKINPTSMYIAPRVS 532
 Db 441 LAAPQLQFAYDKLRASINKVLEELSRACREQVDRDTVMYELSKINPTSMVTAIYGRPV 500
 QY 533 AKRIGDVISNCLVVDQTSVLSHLSRLSASDEKCFSPPTVTFKPMNDSTYKQGLGV 592
 Db 501 AKFVGDAISVDCVAVDQASVSIHKSRL--TSTPGICYSRPPVTFRLNLTTLFKGOLGP 558
 QY 593 NNEILLTTTLETQENTYFYQAKTDMYIKNVEHLKTVPLSSITLDTFIALLPTLLE 652
 Db 559 RNEILLDNQVEACKETCEHYFTASNVTYKDFVFKKINTSEISTLGTFIALLNLSFIE 618
 QY 653 NVDFKVIELTRDEKRLS--NVFDIETMFREYNYAQRVGLRKDL---LDLSTNRNQFVD 708
 Db 619 NIDFRVIELYSRAEKKLSGVSFDIETMFREYNYQTLQAGLREDLONTIDL--NRDLAR 676
 QY 709 AFGLMDDLGAVGQTVVNVASGVATLFPSSIVTGFINKNPPFGGMLMIIVIGVLPFIAYF 768
 Db 677 DLSEIVADLGDVGRTVVNVASSVITLFGSIVSGFINFIKSPFGGMLMILVIVAVLIVFA 736
 QY 769 LTKKTKIYETAPIKMIYPEIDKLEREGKSEIAPISEEELERIVLWMIHQSHMETKT 828
 Db 737 LNRKNTAIAQAPRMVYDIDKQPSGGK-----VDOEQIKNLAGHQLQGBERRLDE 791
 QY 829 RKDPKDSILTRAQNMRLKR--SGYSLNKNABSEVE 860
 Db 792 QORSAPSLFRASDGLKRRRGYKPLENEBAQE 824

RESULT 4
 ID AAU14094 standard; peptide; 856 AA.
 AC AAU14094;
 AC AAU14094;
 XX 21-NOV-2001 (first entry)
 XX Peptide sequence from Epstein-Barr virus glycoprotein precursor gp115.
 DE Anti-retroviral; DP178-like; DP107-like; antifusogenic; antiviral;
 KW Glycoprotein gp110 precursor; gp115; HIV transmission.
 XX Epstein-Barr virus strain B95-8.
 XX WO200151673-A2.
 XX 19-JUL-2001.
 XX 05-JUL-2000; 2000WO-US835727.
 XX 09-JUL-1999; 99US-0350841.
 XX (TRIM-) TRIMERIS INC.
 PA Jeffs P, Lackey JW, Erickson JB, Lawless MK, Merutka G;
 PI WPI; 2001-442157/47.
 DR Identifying a compound that inhibits the formation of or disrupts a
 PT DP107/DP178 complex, especially compounds with antifusogenic, antiviral
 PT or intracellular modulatory activity, by detecting the formation of a
 PT DP107/DP178 complex -
 XX Disclosure; Fig 32; 259pp; English.
 PS The present invention relates to peptides which exhibit anti-retroviral
 CC activity. The peptides of the invention (AAU12559-AAU14009) comprise
 CC DP178-like' and DP107-like peptides. The DP178 peptide corresponds

CC to amino acids 639-673 of the transmembrane protein gp41 from human
 CC immunodeficiency virus 1 (HIV-1) isolate LAI. The DP107 peptide
 CC corresponds to amino acids 558-595 of gp41 from HIV-1LAI. The invention
 CC also relates to a method of identifying compounds that inhibit the
 CC formation of or disrupts a DP107/DP178 complex. The method comprises
 CC detecting the formation of a DP107/DP178 complex, both in the presence
 CC or absence of a test compound, in a reaction mixture containing DP107
 CC and DP178 peptides. The method is useful for identifying compounds,
 CC including small molecule compounds, which may themselves exhibit
 CC antifusogenic, antiviral or intracellular modulatory activity. The
 CC DP178-like/DP107-like peptides are useful to inhibit human and non-human
 CC retroviral, particularly HIV, transmission to uninfected cells. The
 CC present sequence represents a peptide sequence from Epstein-Barr virus
 CC strain B95-8 glycoprotein gp110 precursor, gp115.
 XX
 SQ Sequence 856 AA;
 Query Match 40.5%; Score 1840; DB 22; Length 856;
 Best Local Similarity 45.8%; Pred. No. 7.1e-127;
 Matches 370; Conservative 142; Mismatches 234; Indels 62; Gaps 13;
 QY 81 STFPYRVCASGVGDVFRFQTDHVCDDASDMV-HSEGILLIYKQNIIPMFRVRKRVKV 139
 Db 44 TSFPFRVCELSSHGDLFRSSDIQCPSGCTRENHTEGLLMVFKDNIIPYSFKVRSYTKIV 103
 QY 140 TTSTVYNGIYSDITNQHFTFYKSIPEWTEKWDITTYOCFNSLRNLNTGGNLLTYVDRDDIN 199
 Db 104 TILIYNGIYADVSNVNRHEEKFSVDSYDQDMDTIYQCNVAVMTKDGITRVVYVDRDGN 163
 QY 200 MTVFLOQVDEGVTDPVKRYGSGQPELYLEPGWFGSVRRRTTVNCELMDMFARNSPPDFV 259
 Db 164 IIVNLKPTGGLANGVRRYASQTELYDAPGMLIWTYTRITVNCILITDMMAKNSPDDFFV 223
 QY 260 TATGDTVEMSPWGSDEDDHENKMKHPFVSVINNVYVQYONRGTVPIGLKTRIFLDREE 319
 Db 224 TTTGTQTVEMSPFYDGK--NKETFHERADSFHVRTNYKIVDYDNRGTPNQGERRAFLDKGT 281
 QY 320 YTLSEKHLKKN--MSYCPGLTLKAFYNGIOTESHSGSVHFVANDITASFTTSKEDMKFNIT 378
 Db 282 YTLSMK--LENRTAYCPLQHWQTFDSTIATETGKSIHFVDEGTSSFTVNTTVGIELDPA 339
 QY 379 YHCLNEBEIKAEIKKAKYNSTHSHYKGLK-----YFKTDGGLYLVNQPLQIR 427
 Db 340 FKCIIEQ-----VNKTHEYEAQDRTYKQCAITYFITSGGILLAWLPLTPRS 388
 QY 428 LLDANK--KLNNETYS-----RRSRQAESTTDPMMEMTNGA 463
 Db 389 LATVKNLTELTTTSPSPSPAPSAARGSTPAAVLRRRRRRDAGNATTP---VPPTAP 445
 QY 464 GGEYSSENSITVAQVQYAYDNLIRINNILEDLKAWCREQRAALVWNLKSKINPTSMV 523
 Db 446 GKSLGTLLNPATVQIQFAYDSLRRQINRMGLDLARAWCLEQKRONMVLRELTKINPTTV 505
 QY 524 SMYINRPVSAKRIGDVISVSNCLVVDQTSVLSHLSRLSASDEKCFSPPTVTFKPMND 583
 Db 506 SSIYGKAAVAKRLGDVISVSCVPVNOATVTLKSNR--VPGSETMYSRPLVSFSFINDT 564
 QY 584 TTYKGOLGVNNEILLTTTLETQENTYFYQAKTDMYIKNVEHLKTVPLSSITLDTFI 643
 Db 565 KTYEGOLGTDNEIFLTKMTEVCQATSYQYFSGNIEHVYNDYHFKTEILOGIATLQTF 624
 QY 644 IALNFTLLENVDPKVIELYTRDEKRLSNVFDIETMFREYNYAQRVGLRKDL--LSTN 702
 Db 625 ISLNTSLINIDFASLELYSRDEQRASNVDFLEGIFREYNFQANJAGRLKDLNADVNSG 684
 QY 703 RNQFVDAFGLMDDLGAVGQTVVNVASGVATLFPSSIVTGFINKNPPFGGMLMIIVIGV 762
 Db 685 RNQFVDGLGELMDSLGSVQSITNLVSTVGGFLFSSLVSGFISFFKNPFGGMLILVLVAGV 744
 QY 763 LFAIYFLTKTKIYETAPIKMIYPEIDKLEREGKSE---IAPISEEELERIVLWMIHQ 819
 Db 745 VILVILSTRTRRQMSQOPVQLYPGIDELAQOQHASGEGGFINPISKTELQAIMLA--LHE 802

PT antibodies, useful in vaccines
XX Example; Fig 3; 58pp; English.

CC The SQ in AAN70455 is in the opposite orientation to the prototype
CC orientation of the HCMV genome. The inventors claim antigenic
CC determinants which correspond to a part of the protein coded by the
CC HindIII F fragment of the HCMV genome (between 1378 and 4095 bases
CC from the F/D boundary) and/or the HindIII L fragment (228-2456
CC bases from the L/D boundary). The antigens are useful in vaccines
CC against HCMV and for detecting antibodies specific for HCMV.

XX Sequence 906 AA;

Query Match 24.7%; Score 1121.5; DB 8; Length 906;
Best Local Similarity 30.4%; Pred. No. 8.3e-74;
Matches 282; Conservative 186; Mismatches 350; Indels 111; Gaps 25;

QY 14 LWLYQVALYSLSTAETGVTSPPTATWSTESPLTGHGTHDSHGGRGN-ENRD--- 68
DB 5 IWCL--VVCNLCIVCLGAUVSSSTSHATSS-----THNGSHTSRRTSAQTRSVSQ 55
QY 69 -----SEQONKIYG-----SPSTFPYRVCSASGVDFRFQTDHVC--- 105
DB 56 HVTSEASVSHRANETIYNTLLKYGDVGVNTKYPRVCSMAQGTDLIRFERNICTSMK 115
QY 106 PDASDMVHSEGILLIYKONIIPFMRVRKYRKVTTSTVNGIYSDSITNQHTFYKSI 165
DB 116 PINEDL--DEGIMVYKRNIVAHTEKRVYQKVLTFRRSYAYIYTYLLGSNTYEYVAPM 173
QY 166 WETEKMDTIYOCFNSRLNTGCLLTIVDRDI--NMTVFLQPDVGVTPDVKRYGSOPELY 224
DB 174 WEIHHINKFAQCYSSYRVIGTGVFVAYHRDSYENKTWQLIPDDYSNTHSTRIVTVKQDW 233
QY 225 LEPGFWGYSRRRTTIVNCELDMFARSNPPDFEVTATGDTVEMSPFWSGEDDHENKHE 284
DB 234 HSRGSTW-LYRETCLNCLMTITARRSKYPVHFATSTGDVVYISPFYNGTRNASYFGE 292
QY 285 KPWFEVSNVNNYKV-VQVNGRTVP-LGKTRIFLDREYTLSE-KHLKNMSYCPILTWKA 341
DB 293 NADKFFIPPNYTVISDFGRPNAAPEHRLVAFLEADSVISWDIQDEKNVT-CQLTFWEA 351
QY 342 FYNGIQTEHSGSYHFVANDITASPTTSKEDMKENTTYHCLNEIKAEIKKY-AKVNST 400
DB 352 SERTIRSEADSYPHSAKMTATFLSKQEVNMSDSALDCVRDEAINKLOIENWYNQT 411
QY 401 HSKYGDLYKFTDGLYLVWQPLQLNRLDAKKNLNNTYS---RRSRQAESTTDPME 457
DB 412 YEKYGNVSVPETSGLVVFWQGIKQKSLVELERLANRSSLNITHRTTRSTSDNNTLSS 471
QY 458 MTGNAGAGEYSSENSITVAQVQAYDNLRIINNILEDLSKAMCREQRAALVWNLSKI 517
DB 472 M-----ESVHNLYVAQLQFTYDTRLGYINRALAQIAEAWCVDQRRTLVEFKELSKI 522
QY 518 NPTSVMSIMNRPVSARKIGDVTISVNCIVVDOTSVSLHKLRLLSASDEKCFSPRPVTF 577
DB 523 NPSAILSAINKPIAPAFMGDVLGACVINTQTSVKVLKDMN-KVESPGRCISRPVPIF 581
QY 578 KFMNDSTIYKQGLGVNNEILLTTTYLETQCENTEYFYQAKTDMYIYKNYEHKLTVP LSSI 637
DB 582 NFANSSVVOVQGLGEDNEILLGNHRTBECQPSLKIFAGNSAYEYVDYLFKRMIDLSSI 641
QY 638 TLLDTFTALNFTLLENVDFKVIELYTRDEKLSNVDFIETMFPREYNYAORVSGRLKLL 697
DB 642 STVDSMALDIDPLENTDFRVELYSQKELSSNVDFLEIEIMREFNSYKORVRYVEDKV 701
QY 698 DLSTNRNQVDAFGSLMDDLCAVGQTVVNAVSGVATLFFSIVTGFINFINKNPFGLMII 757
DB 702 D---PLPPYKLGDLDDLSGGLGAGKAVGVAIGVAGGAVASVGEVATFLKNPFGATILL 758
QY 758 VVIGLVFAIYFLTKTKIYETAPIKMIYPEI-----DKLKE----- 793
DB 759 VAIYAVIITLIYTRORRLCTQPLQNLFPYLVNSADGTTVTSGTKDTSLOAPPSEESVY 818

QY 794 ---REG-----KSEIAPISEEELERIVLAM-----HIHQONS-----HMETKTR-KD 831
DB 819 NSRGKGGPSSDASTAAPPTNEQAYQMLLALRLDABQRAQQNGTSLDGGTQDQKG 878
QY 832 PKDSILTRAQNLMLRKSGYSNLKNAESVE 860
DB 879 QKPNLLDLRLH---RKNGYRHLKDSDEEE 904

RESULT 9
AAP91047

ID AAP91047 standard; protein; 906 AA.

XX AAP91047;

DT 04-JAN-1990 (first entry)

XX gB envelope protein of cytomegalovirus.

XX Cytomegalovirus; glycoprotein gp. 55; glycoprotein gB;
XX vaccine; prophylactic agent.

XX Cytomegalovirus.

XX WO8907143-A.

XX 10-AUG-1989.

XX 26-JAN-1989; 89WO-US00323.

XX 29-JAN-1988; 88US-0149715.

XX (CHIR) CHIRON CORPORATION.

XX Spaete RR, Pacht CA;

XX WPI; 1989-249012/34.

XX New recombinant polypeptide from gB gene of cytomegalovirus - useful in
XX vaccines, also diagnostic DNA coding sequences and derived antibodies.

XX Claim No. 1; fig No. 2; 64pp; English.

XX The gB envelope protein derived from glycoprotein gp. 55 contains an
XX epitope immunologically identical to one encoded by the CMV genome, and
XX is useful in vaccines and as a prophylactic agent against CMV
XX infections, or for detecting CMV and CMV Abs. See also AAN90528.

XX Sequence 906 AA;

Query Match 24.5%; Score 1114; DB 10; Length 906;

Best Local Similarity 30.7%; Pred. No. 3e-73;

Matches 284; Conservative 184; Mismatches 355; Indels 102; Gaps 25;

QY 14 LWLYQVALYSLSTAETG--VTSPTATWSTESPLTGHGTHDS--HGERGNENRDS 69

DB 5 IWCL--VVCNLCIVCLGAUVSSSTTRGTSATHS---HHSHHTTSAHSRSGSVQRVT 58

QY 70 EEQ-----NKNIYG-----SPSTFPYRVCSASGVDFRFQTDHVC---PDA 108

DB 59 SSQTVSHGVNETIYNTLLKYGDVGVNTKYPRVCSMAQGTDLIRFERNICTSMKPIN 118

QY 109 SDMVHSEGILLIYKONIIPFMRVRKYRKVTTSTVNGIYSDSITNQHTFYKSI 168

DB 119 EDL--DEGIMVYKRNIVAHTEKRVYQKVLTFRRSYAYIYTYLLGSNTYEYVAPMWEI 176

QY 169 EKMDIYOCFNSRLNTGCLLTIVDRDI--NMTVFLQPDVGVTPDVKRYGSOPELYLEP 227

DB 177 HHNGHSQCYSYSRVYAGTVFVAYHRDSYENKTWQLMPDDYSNTHSTRIVTVKQDWHR 236

QY 228 GWFWSYRRRTTVNCELMDMFARSNPPDFVTATGDTVEMSPFWSGEDDHENKHEKWP 287

```

Db 237 RSTW-LYRETCLNLCMVTTTARSKYPHFFATSTGDDVVDISPFYNGTNRNRSYFCENAD 295
Qy 288 FVSINNYKVVDYQNRGTVPGLGKTRI--FLDREETYLSWE-KHLKMSYCPLTLWKAFYN 344
Db 296 KFFIFPNYTVISDFGRPNSEALETHRLVAFLEADSVISWDIODEKVT-CQJTFWEASER 354
Qy 345 GQTEHSGSYHFVANDITASFTTSKEDMKFNTTYHCLNEERIKABIEKKY-AKVNSTHSHK 403
Db 355 TIRSEADSYHPSAKMTATFLSKQEVNMSDALDCVRDEAINKLQIFNTSYNQTYEK 414
Qy 404 YGDLKXFKTDGGLYLVWQPLQNRLLDAKLNNETYSRSPROAESTITDPMENTGGA 463
Db 415 YGNVSUFETGGLVFWQIQKSLVELERLANRSLNTHRTKRST-----DGNNA 467
Qy 464 G--GEYSSENSITVAQOVAYONLRIRINNILEDLSKAWCREQRAALVWNELSKINPTS 521
Db 468 THLSNMESVHNLVYAQLQFTYDTRAGYINRALAQAEAWCVDQRTLEVFKLSKINPSA 527
Qy 522 VMSMIYNRVPSAKRIGDVISVNCIVVDQTSVLSHLKSLRLLSASDEKCFSPRPVTFKFN 581
Db 528 ILSAIYNKPIAARFMGDVLGLASCVTINQTSVKVLRDMN-VKESPGRCYSRPPVWFNFAN 586
Qy 582 DSTIYKGLGVNNEILLTTLTYLETQENTYEFQAKTDMYIYKNYEHKTVPLSSITLID 641
Db 587 SSYVQVGLGEDNEILLGNHRTBECOLPSLKIFAGNSAYEYVDYLFKRMIDLSSISTVD 646
Qy 642 TFIALNFTLLENVDFKVIELYTRDEKRLSNVFDIETMFREYNYAQRVSGLRKDLDLST 701
Db 647 SMALDIDPLENTDFRVLALYSQELRSNVDLEIEIMREFNSYKQRYVEDKVD--- 703
Qy 702 NRNQVDFAPGMLDGLGAVGTQVNAVSGVATLFSISVTGFINFKNPGGMLMIIVVIG 761
Db 704 PLPPYLKGLDLDLMSGLGAAGKAVGAIGAVGAVSVEGVATFLKNPFGAFTIILVAIA 763
Qy 762 VLFAYFLTKTKIYETAPIKMIYPEI-----DKLKE-----R 794
Db 764 VIIILYMTQRRLCMOPLQNLFFYLVASADGTTVTSGNTKDTSLQAPPSYEESVYNSGR 823
Qy 795 EG-----KSEIAPISEELEERIVLAM-----HIHQONS---HMETKTR-KDPKDS 835
Db 824 KGPSPSSDASTAAPPYNEQAYQMLLALVRLDAEQRAQNGTDSLQGTQDQKQKPN 883
Qy 836 ILTRAQNMLRKSGSYNLSKNAESVE 860
Db 884 LLDRLRH---RKNGYRHLKDSDEE 905

RESULT 10
AAB31678
XX AAB31678 standard; Protein; 195 AA.
AC AAB31678;
XX
XX 30-APR-2001 (first entry)
DE A viral glycoprotein B envelope protein fragment.
KW Glycoprotein B; envelope protein; vaccine; passive immunity; pig.
XX Swine gamma herpesvirus.
OS WO200103733-A1.
PN 18-JAN-2001.
XX 07-JUL-2000; 2000WO-US18647.
XX 08-JUL-1999; 99US-0142736.
XX 02-DEC-1999; 99US-0168532.
XX (BIOT-) BIOTRANSPLANT INC.
XX Patience C;
PI

```

```

XX WPI; 2001-103033/11.
DR N-PSDB; AAF25178.
XX
XX New polynucleotide encoding a polypeptide corresponding to porcine
PT gamma-herpesvirus glycoprotein B (pGHV-gpB) for immunizing a pig
PT against a porcine gamma-herpesvirus -
XX
XX Disclosure; Page 67; 82pp; English.
XX
XX The present sequence represents a fragment of a swine gamma herpesvirus
CC glycoprotein B envelope protein. Probes and primers derived from DNA
CC encoding the glycoprotein B envelope protein are used for detecting the
CC presence of swine gamma herpesvirus. Antibodies directed against the
CC -envelope protein are also useful for detecting presence of the virus.
CC The envelope protein can be used as a vaccine for creating passive
CC immunity in a pig.
XX
XX Sequence 195 AA;
SQ
Query Match 22.0%; Score 1002; DB 22; Length 195;
Best Local Similarity 100.0%; Pred. No. 5.8e-66;
Matches 195; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 484 NLRIRINNILEDLSKAWCREQRAALVWNELSKINPTSVMGMIYNRPVSAKRIGDVISVS 543
Db 1 NLRIRINNILEDLSKAWCREQRAALVWNELSKINPTSVMGMIYNRPVSAKRIGDVISVS 60
Qy 544 NCIVVDQTSVLSHLKSLRLLSASDEKCFSPRPVTFKFMNDSTIYKGLGVNNEILLTTL 603
Db 61 NCIVVDQTSVLSHLKSLRLLSASDEKCFSPRPVTFKFMNDSTIYKGLGVNNEILLTTL 120
Qy 604 ETCQENTYEFQAKTDMYIYKNYEHKTVPLSSITLDTFIALNFTLLENVDFKVIELYT 663
Db 121 ETCQENTYEFQAKTDMYIYKNYEHKTVPLSSITLDTFIALNFTLLENVDFKVIELYT 180
Qy 664 RDEKRLSNVFDIETM 678
Db 181 RDEKRLSNVFDIETM 195

RESULT 11
AAR77024
ID AAR77024 standard; Protein; 879 AA.
XX
XX AAR77024;
AC AAR77024;
XX
XX 20-APR-1996 (first entry)
DE Canine herpesvirus gB homologue.
KW CHV; glycoprotein gB; vector; attenuation; poxvirus; vaccinia virus;
KW canarypox virus; ALVAC; vCP320; dog; puppy; immunisation; antigen;
KW vaccine.
XX
XX Canine herpesvirus.
OS
XX
XX Key Location/Qualifiers
FH Region 725..741
FT /label= Transmembrane_region
FT Region 747..771
FT /label= Transmembrane_region
XX
XX WO9526751-A1.
PN
XX 12-OCT-1995.
XX
XX 30-MAR-1995; 95WO-US03982.
XX
XX 29-MAR-1995; 95US-0413118.
XX 30-MAR-1994; 94US-0220151.
XX
XX (VIRO-) VIROGENETICS CORP.
PA

```


Db 744 LHLREYDI--DRVKTGDMNMRGLANFPOGLGAVGQAVGTIVLGAAGAALSTVSGTA 801
 QY 744 NFIKNPFGMLMIIVVIGVLFYFKTKIVETAPIKIWIYPIDKLKEREGKSEIAP 803
 Db 802 SFIANPFGALATGLVLGVAFLAYRYSIRLSNPMKALPITTRALKDDARGATAPG 861
 QY 804 SEEE-----LER---IVLAMIHQONSHMETKTRKDPKDS--ILTR-AQNMLRKRS--G 849
 Db 862 EEEEFDAKLEQAREMIKYMSLVSAVERQEHKAKSKNGGKGPLLATRLTQLALRRAPPE 921
 QY 850 YSNLKNAE 857
 Db 922 YQOLPMAD 929

RESULT 15

AAR14665
 ID AAR14665 standard; Protein; 904 AA.

XX AAR14665;

XX AC

DT 25-MAR-2003 (updated)

DT 23-JAN-1992 (first entry)

XX HSVB polypeptide.

DE Vaccine; antigen.

XX Herpes simplex virus.

OS JP03218397-A.

XX 25-SEP-1991.

XX 21-JUN-1990; 90JP-0161448.

XX 30-NOV-1989; 89JP-0308941.

PR 22-JUN-1989; 89JP-0158238.

PR 21-JUN-1990; 90JP-0161448.

XX (TAKA) TAKEDA CHEM IND LTD.

XX WPI; 1991-328397/45.

DR N-PSDB; AAQ14478.

XX HSVB polypeptide(s) obt'd. by recombinant DNA techniques -

PT useful as vaccines against HSV and in diagnosis, can be produced

PT cheaply and safely.

XX Claim 1; Fig 2; 24pp; Japanese.

XX The sequence was deduced from DNA and is that of the HSVB polypep-

CC tide. The recombinant protein or esp. truncated forms of it, can be

CC used to prepare vaccines for prophylaxis of HSV infection and for use

CC in diagnostic kits.

CC See also AAR14666.

CC (Updated on 25-MAR-2003 to correct PA field.)

XX Sequence 904 AA;

SQ

Query Match 18.6%; Score 847.5; DB 12; Length 904;

Best Local Similarity 27.7%; Pred. No. 1.4e-53;

Matches 244; Conservative 180; Mismatches 362; Indels 95; Gaps 27;

QY 28 APTGVTSPNNTATWSTESP--LTGHYGTDSHGE--RGNENRDESEQNKNTYGSPTFP 84

Db 62 APTGDTKPKKKNKPPPPRAGDNATVAAGHATLREHLRDIKAKNTDANFY----- 114

QY 85 YVCSASGVDFRQTDHVCPSDSD--MVHSGLILLIKQNIIPFMRVRKYRKVVTTST 143

Db 115 --VCPPTGATVQVEQPRRCPTREGQNYTEGIAWFKENIAPYKFKATMYKDVTSVQ 172

QY 144 V-----YNGIYSDSITNQHTFYKSIPEWE--TEKMDTIYQCFNSRLNTGNNLLTYV 193
 Db 173 VMFGHRYSQFNGIIFED-----RAPVPEEVIDKINAKGVCRSTAKYVRRNNLETTAF 223
 QY 194 DRDDINMTVFLOPDGV-----TPDVKYGSOPELYLEPGHFWGYSRRRTTVNCML 245
 Db 224 HRDHETDMELKPANAAATRTSRGWHTTDLKYNPSRVEAF-----HRYGTTVNCIVE 274
 QY 246 DMFARSNPFPFFVATGDTVEMSPFWS--GEDDH--ENKMHKPFVSVINNY--KVVDYQN 302
 Db 275 EVDARSVPYNEFVLATGDFVMSPPFYCYREGSTEHTSYAADRFKQVDGFYARDLTTKA 334
 QY 303 RGTPLGKTRIFLDREYYTLSWEKHLKNMSPCLTLMKAFYNGIQTEHSGSYHFVANDIT 362
 Db 335 RATAP--TTRNLLTTPKFTVAMDVPKRPSTCTWKQEVDEMLRSEYGGSFSSDAIS 392
 QY 363 ASFTT-----SKEDMKFNTTTHCLNEETKAEIKKYA--KVNSTHSKYGDLKYFKTDG 414
 Db 393 TTTTNTLTPSLRVDLGD-----CIGKDARDAMDRIFARRYNATHIKVGPQYYLANG 446
 QY 415 GLYLWQPLIQNRLLDAKNLNNETYSRRSROAESTTDPMMEMTNGCAGGEYSSENSIT 474
 Db 447 GLIAYQPLLSNTLAE-----LYVREHLREQRKPNPTPPPGASANASVERIKTTSSIE 502
 QY 475 VAQVQAYDNLRIRINILEDLSKAWCREQHRALVMNELSKINPTSYMSIYNRPVSAK 534
 Db 503 FARLQFTYNNHQRHVNDMLGRVAJAWCFELQNHLELTWNEARKLNPNATASVTGRRVSAR 562
 QY 535 RIGDIVSNCIVVDQTSVSLHLSASDEKCFRSPPTVTFKPMNDSTIYKQOLGVNN 594
 Db 563 MLGDVMAVSTCPVVAADNVIQNSMR--ISSRPGACYSRPLVSFRYEDQGPLVEGOLGNN 621
 QY 595 EILLTTTLETQENTTEYFOAKTDMYIKNVEHLKTVPLSSITTLDTFIALNFTLLENV 654
 Db 622 ELRTRDAIEPCTVGHRRYFTFGGYVYFEEYAYSHQLSRADITTVSTFDLNTIMLEDH 681
 QY 655 DFKVIELYTRDEKLSNVFDIETMFREYNYVAQVSGRLKDLDLSTNRNOF--VDAFGS 712
 Db 682 EFVPLEVYTRHEIKDSGLDYTEVQRRNQLHDLRFADI--DTVIHADANAAMFAGLGAFFE 740
 QY 713 LMDDLGLAVGQTVNAVSGVATLFSIVTGFINIKPFPFGMLMIIVIGVLFALYFLTK 771
 Db 741 GMGDLGRAVGKVMGIVGGVV-----SAVSGVSSFSMSPFGALAVGLLVLAGLAAAFAPR 796
 QY 772 KTKIYETAPIKIWIYPIEDK-LK-----EREKGSIAPISEB-----ELERIVLAMHIHQON 821
 Db 797 YVMLQSNPMKALYPLTTKELKNTNPDASGEGEGGDFDEAKLAAREMIRYMALVSAM 856
 QY 822 SHMETKTRKDPKDSILT-RAQNML---RKRSGYNLKNAES 858
 Db 857 ETEHKAKKKGTSALLSAKVTDVMVRKRNTNYTQVPNKDS 897

RESULT 16

AAP71135

ID AAP71135 standard; Protein; 903 AA.

XX AAP71135;

XX AC

XX 25-MAR-2003 (updated)

DT 30-APR-1991 (first entry)

XX Herpes Simplex Virus-1 GB from pKEXX.

XX Vaccine; prophylaxis; treatment; Herpes Simplex Virus-1;

XX glycoprotein; GB.

XX Herpes simplex virus type 1 (KOS).

XX US4642333-A.

PN 10-FEB-1987.

XX

```

PF 20-JUN-1984; 84US-0622496.
XX
XX
XX 20-JUN-1984; 84US-0622496.
XX (PERS/) PERSON S.
XX Person S;
XX WPI; 1987-056354/08.
XX N-PSDB; AAN71303.
XX
XX Amino acid chain of glycoprotein B of HSV-1 and 2 - prepd. as
XX recombinant and used for vaccines for herpes simplex virus types 1
XX and 2.
XX
XX Example; Table 2; 16pp; English.
XX
XX A pure non-glycosylated amino acid (AA) chain comprising a sequence
XX corresponding to that occurring in glycoprotein B of HSV-1 or HSV-2
XX virus which is antigenic to HSV-1 of HSV-2, which contains no more
XX than 750 AA residues, and which includes AA residues 135-649
XX inclusive is claimed. It can be used to produce vaccines for
XX prophylaxis and treatment of HSV-1 and HSV-2.
XX See AAN71399 for the HSV-2 sequence.
XX (Updated on 25-MAR-2003 to correct PR field.)
XX
XX Sequence 903 AA;
XX
XX Query Match 18.6%; Score 845.5; DB 8; Length 903;
XX Best Local Similarity 27.6%; Pred. No. 1.9e-53;
XX Matches 245; Conservative 181; Mismatches 360; Indels 101; Gaps 28;
XX
XX 24 SLSIAETGVTSPPTATWSTSP--LTHGYTHDSHGE--RGNNENRDSBEQNKVIYGP 80
XX 57 ALGAAPTGDPKPKXKPKNPPTPRPADGNATVAAGHATLREHURDIKAENTDANFY--- 113
XX 81 STEFVRVCSAGVGDFRQFDHVCPSD--MVHSEGLLIYKONIIIPFMRVRKRVKV 139
XX 114 -----VCPPTGATVQFQPRCPTPEQNTGIAVVFKENIAPYKFKATMYKDV 167
XX 140 TTSTV-----YNGIYSOSITNQHTFYKSIPEW--TERMDTIYOCFNSLRNTGNNL 189
XX 168 TVSQWFGHRYSPQWGFIED-----RAPVPFEEVIDKINAKGVCSTAKYVRNNLE 218
XX 190 LTYVDRDINMTVFLQPDGV-----TPDVKRYGSOPELYLEPGFWGVSRRRTTN 241
XX 219 TTAHFRDDHETDMLKANAATRSRGWHTTDLKYNPSRVEAF-----HRYGTVN 269
XX 242 CELDMFARNPPDFVTATGDTVMSPFWS--CEDDH--ENKHEKPFVSVINNY--KV 298
XX 270 CIVEVDARSVYDFEVLATGDFVNSPFFYGYREGSHTHTTYAADRFKQVDGFYARDL 329
XX 299 DYQNRGTVPGLKTRIFLDREYTLSEWKLKNSYCPLTLWKAFYNGIOTHEGSGYHFA 358
XX 330 TTKARATAP--TTRNLLTPKFTVANDWVPRKPSVCTMTKWQEVDEMLRSEYGSFRSS 387
XX 359 NDITASFTT-----SKDMKENTYTHCLNEIEKAEIKYA--KVNSTHSHKVGDLKYP 410
XX 388 DAISTTTTNLTNTEPLSRVDLGD-----CIGKDARDMRIFARRYNATHIKVGPQYY 441
XX 411 KTDGGLVLMVOPLQLNLLDAKNLNNETYSRRSRROAESTTDPDMEMTNGAGGESYE 470
XX 442 LANGFLLIAQPLSLNLAEE-----LYVREHREOSRKPPTPPPPGASANASVERIKTT 497
XX 471 NSITVAQVQAYDNLRIINNIIDLSKAMCREQRAALVWNLKINPTSVMSMIYNRP 530
XX 498 SSIEFARLQFTYNIQRHVMNMLGRVAIAWCELQNHETLWNEARKLNPNNAIASVTVGR 557
XX 531 VSAKRIGDVISVNCIVVDQTSVLSHLKSLLSLSDSKSRPPTVKFNDSTIYKQL 590
XX 558 VSARMLGDVNAVSTCVPAADNVIVQNSMR--ISSRPGACYSRPLVSPRYEDQGPLVEGOL 616
XX 591 GVNNEILLTLYETCQENTEEYFQAKTMDMYIKVYEHKLTVPPLSSITLDTFIALNFTL 650
XX
XX Db 617 GENNELRLTRDAIEPCTVGHRRYFTFGGYYVEEYAYSHQLSRADITTVSTFDLNI 676
XX Qy 651 LENVDFKVIETYTRDEKRLSNVFDIETMFREYNYAYQORVSGRLKDLJLSTNRNQF--VD 708
XX Db 677 LEDHEFVPLEVYTRHEIKDSGLLDYTEVQRNQLHLRFADI-DTVIHADANAAMFAGLG 735
XX Qy 709 AFGSLMDDLG-AVGQTVNAVSGVATLFPSSIVTGFINFKNPFGGMLMIIVIGLFAIY 767
XX Db 736 AFFEGMDLGRVGVKVMGIVGVW----SAVGSVSFMSNPPGALAVGLVLGLAAAF 791
XX Qy 768 FLTKTKIYETAPIKMIYPEIDK-LK-----EREKKEIAPISE-BELERIVLA 814
XX Db 792 FAFRYVMRLQSNPMKALYPLTTKELKMPNPDASGESEGGDFDEAKLAAREMIRYMAL 851
XX Qy 815 MHIHQNSHMETKTRKDPKDSILT-RAQNMWML-----RKRSGYSNKNNAE 857
XX Db 852 VSAMERTEH---KAKKGTSLLSAKVTDMVKRRRNTNYTQVFNKD 895
XX
XX RESULT 17
XX AAE17812
XX ID AAE17812 standard; Protein; 904 AA.
XX AC AAE17812;
XX XX
XX DT 01-MAY-2002 (first entry)
XX XX
XX Herpes simplex virus HSV-2 UL27 antigenic protein.
XX DE
XX Herpes simplex virus; HSV antigen; therapy; human immunodeficiency virus;
XX KW HIV infection; vaccine; virucide; HSV-2 UL27 protein.
XX XX
XX Herpes simplex virus.
XX OS
XX WO200202131-A2.
XX PN
XX PD 10-JAN-2002.
XX XX
XX PF 28-JUN-2001; 2001WO-US20981.
XX XX
XX PR 29-JUN-2000; 2000US-215458P.
XX PR 20-MAR-2001; 2001US-277438P.
XX XX
XX PA (CORI-) CORIXA CORP.
XX XX
XX PI Hosken NA, Day CH, Dillon DC, McGowan P, Sleath PR;
XX XX
XX WPI; 2002-154689/20.
XX DR
XX PT Novel isolated polypeptide comprising at least an immunogenic portion
XX of herpes simplex virus antigen, useful as component of vaccines used
XX for treating herpes simplex virus infection in a patient
XX PS
XX Claim 1; Page 125-127; 157pp; English.
XX XX
XX CC The patent discloses polypeptides comprising Herpes simplex virus (HSV)
XX antigens, DNAs encoding HSV antigens and the use of such compositions
XX and methods for the diagnosis and treatment of HSV infection. The
XX compounds comprise polypeptides that contain at least one immunogenic
XX portion of an HSV antigen. Proteins of the invention are useful for
XX detecting human immunodeficiency virus (HIV) infection in a patient
XX and for treating HSV infection. They are also useful as component of
XX vaccines used for treating herpes simplex virus infection in a patient.
XX CC The present sequence is Herpes simplex virus HSV-2 UL27 antigenic
XX protein of the invention.
XX XX
XX SQ Sequence 904 AA;
XX
XX Query Match 18.6%; Score 844; DB 23; Length 904;
XX Best Local Similarity 28.4%; Pred. No. 2.5e-53;
XX Matches 247; Conservative 164; Mismatches 366; Indels 92; Gaps 27;

```



```

Db      806 MKALYPLTTKELKUNTPNDPDSGEGEGDFDEAKLAAREMIRYMALVSAMERTEH---K 862
Qy      828 TRKDPKDSILT-RAQNML---RRKSGYSLNKNAE 857
Db      863 AKKKGTSALLSAKVTDVMRKRRNTNTYQVPMKD 896

RESULT 20
ID      AAB74441
XX      AAB74441 standard; protein; 904 AA.
AC      AAB74441;
XX      29-MAY-2001 (first entry)
XX      Herpes simplex virus 1 glycoprotein B.
XX      HSV-1; HSV-2; glycoprotein B; gB; transmembrane envelope glycoprotein;
XX      antigenic epitope; diagnosis; vaccine.
XX      Herpes simplex virus type 1.
XX      US6197497-B1.
XX      06-MAR-2001.
XX      19-APR-1996; 96US-0632537.
XX      21-APR-1995; 95US-0426604.
XX      (UYNE-) UNIV NEW MEXICO STATE.
XX      Goade DE, Bell R, Jenison S;
XX      WPI; 2001-256360/26.
XX      Continuous, isolated, antigenic polypeptide segment of herpes simplex
XX      virus (HSV) glycoprotein B1 or B2, useful in serodiagnostic
XX      immunoassays for distinguishing HSV-1 infection from HSV-2 in a human -
XX      Claim 8; Fig 4; 23pp; English.
XX      The present invention provides antigenic peptides from herpes simplex
XX      virus type 1 (HSV1) and 2 (HSV2) glycoprotein B (gB) which can be used in
XX      the diagnosis of HSV infection, and identification of subtype, and in
XX      CC vaccines to protect against HSV. The present sequence is the HSV-1 gB
XX      protein.
XX      Note: The present sequence is mentioned in claim 8 as being the HSV
XX      glycoprotein B1.
XX      SQ      Sequence 904 AA;

Query Match      18.5%; Score 841; DB 22; Length 904;
Best Local Similarity 27.7%; Pred. No. 4.2e-53;
Matches 242; Conservative 179; Mismatches 345; Indels 108; Gaps 28;

Qy      35 PNTATWSTESPLTCHYTHDSHGE-RGNENRDSBQNKNIYGPSFFPRVCSAGV 93
Db      80 PPRPA-----GDNATVAAGHATLREHLDIKAENTDANFY-----VCPPTG 121
Qy      94 GDVFRFQTDHVCDDASD-MVHSEGLILYKONIIPFMRVRKRYKVTTSTV----- 144
Db      122 ATTVQEQPRCPTRPEQONYTEGIAVFKENIAPYKFKATMYKDVTVSQWFGHRSQ 181
Qy      145 YNGIYSDSITNOHTYKSIKPEW--TEKMDTIYQCFNSLRNLNLTGNNLTIVDRDDINMTV 202
Db      182 FMGIFED-----RAPVPEEVIDKINAKGVCRSTAKYVRNLETTFARDDHETDM 232
Qy      203 FLQPVGCV-----TPDVKEYGQPELYLPGHFWGSGYRRRTVNCELMDMFARSNPP 254
Db      233 ELKPANAARTSRGWHITDLKNPGRVEAF-----HRYGTTVNCIVEEDARSYYP 283
Qy      255 FDFFTATGDTVMSPFFWS-GBDDH-ENKMKHPWFSVINNY-KVDYQNRGTVPGLGKT 311

```

```

Db      284 YDEFVLATGDFVYMSPPFYREGSHTERTSYAADRFXQVDFYARDLTTKARATAP--TT 341
Qy      312 RIFLDREYTLSEWEKHLKNMSYCPGLTWKAPYNGIOTHSYGFVANDITASFTT----- 367
Db      342 RNLTTTPKFTVAVDWVPRPSVCTMTKQVEVDEMLRSEYGGSRFPSSDAISTTFTTNUTE 401
Qy      368 ---SKEDMKFNTTYHCLNEEIKAEIKKYA-KVNSTHSKYGDLYKFKYFTDGGLYVMQPL 423
Db      402 YPLSRVDLGD-----CIGKDARDMDRI FARRYNATHI KVGQPYLYLANGGLFIAYQPL 455
Qy      424 IONRLLDAXKNKNNETYSRRRQAESTTDDPMEMWTGNGAGEYSSSENSITVAQVQYAYD 483
Db      456 LSNLTAE---LYVREHLREQSRKPPNPPTPPPGASANASVERIKTTSSIEFARLQFTYN 511
Qy      484 NLRIRINNILEDLSKAWCREOHRALVWNELSKINPTSVMSMIYNRPVSARKIGDIVSVS 543
Db      512 HIQRHNDMLGRVAIAWCELQNHBLTLWNEARKLNPAIASATVGRRRVSARMGLGDMVAVS 571
Qy      544 NCIVVDOTSVSLHKSLLLSASDEKCFSRPPVTTFKFMNDSTIYKGLQGVNNEILLTLYL 603
Db      572 TCVPAADNVIVQNSMR-ISRPGACYSRPLVSFRYEDOGPLVGEGLGENNELRLTRDAI 630
Qy      604 ETCQENTEYFQAKTDMYIYKNYHLKTVPLUSSITTLDTFTIALNFTLLENVDFKVIELYT 663
Db      631 EPCVTGHRRYFTFGGYVYFEEYAYSHQLSRADITTVSTFDLNTITMLEDFEFPVLEVYT 690
Qy      664 RDEKLSNVFDIETMFREYNYAQRVSLRKLDDLSTNRNQF--VDRFGSMDDLG-AV 720
Db      691 RHEIKDSGLDYTEVQRNQLHDLRFADI-DTVIHADANAAMFAGLAGAFFEGMGDLGRAV 749
Qy      721 GQTVNAVSGVATLFSISITGFINFIKNPFGGLMIIWIGVLFAYIFLTTKTIYETAP 780
Db      750 GKVMGIVGGV----SAVSGVSSFMSPFGALAVGLLVLAGLAAAPFAPRYVMRLQSNP 805
Qy      781 IKMIYPEIDK-LK-----EREKKSIAPISE-EELERIVLAMHIHQONSHMETK 827
Db      806 MKALYPLTTKELKUNTPNDPDSGEGEGDFDEAKLAAREMIRYMALVSAMERTEH---K 862
Qy      828 TRKDPKDSILT-RAQNML---RRKSGYSLNKNAE 857
Db      863 AKKKGTSALLSAKVTDVMRKRRNTNTYQVPMKD 896

RESULT 21
AAW72113
ID      AAW72113 standard; Protein; 854 AA.
XX      AC      AAW72113;
XX      DT      18-DEC-1998 (first entry)
XX      DE      HSV-2 strain SB5 Contig ID 14 ORF#1 protein.
XX      KW      HSV-2 strain SB5; immunological response induction; therapy;
XX      antiviral identification; viral protein inhibitor.
XX      OS      Herpes simplex virus type 2.
XX      PN      WO9820016-A1.
XX      PD      14-MAY-1998.
XX      PF      31-OCT-1997; 97WO-US20016.
XX      PR      09-JUN-1997; 97US-0049018.
XX      PR      04-NOV-1996; 96US-0030279.
XX      (SMIK ) SMITHKLINE BEECHAM CORP.
XX      PA      Chan JY, Dabrowski-Amaral CE, Delvecchio AM, Dillon SB;
XX      PI      Esser KM, Leary JJ;
XX

```


Db	76	TPPPD-----ANATVAAGHATLR--AHLREIKVENADAQ-----FVVCPPPT	115
Qy	93	VGDVFRFOTDHCPCDASD-MVHSEGIILYKQNIIPFMFRVRKYRKVVTSTV-----	144
Db	116	GATVQFEQPRRCPTREPCQNTYEGIAVFKENIAPYKFKATMYKDVTSQVWFGHRS	175
Qy	145	-YNGIYSDSITNOHTFYKIEPWE--TEKMDTIYQCFNSLRNTGNNLLTYVDRDDNMT	201
Db	176	QFMGIFED-----RAPVFEEVIDKINAKGVCRCSTAKYVRNNMETTAFHRDDHETD	226
Qy	202	VFLQPVGDVTPDKVKGQPELYLPGFWGYSRRRTTVNCELMDFARSNPPDFVTA	261
Db	227	MELKPAK-VATRTSRGWHHTDLYKNPSRVEAFHYGTTVNCIVEEDARSVPYDFEFLA	285
Qy	262	TGDTVEMSPFWG-GEDDH-ENKMHKPEFVSVINNY-KVVDYQNRGTVPGLKTRIFLDRE	318
Db	286	TGDFVMSPFYGYREGSHTHTSYAADRFKQVDFYARDLTTKARATSP--TTRNLLTTP	343
Qy	319	EYTLSEKHLKNNMSYCPPLTWKAFYNGIQTEHSGSYHFVANDITASFTT-----SKED	371
Db	344	KFTVAMDVWPKRPVACTMTKQVEDEMLRAEYGGSFSSDAISTTFTTNLTQYSLSRVD	403
Qy	372	MKEFNTTYHCLNEEIKAEIEKKYA-KVNSTHSKYGDLKFKYKTDGGLYLWQPLIQNRLLD	430
Db	404	LGD-----CIGRDAREALDRMFARKYNATHIKVGQPOYYLATGGLFIAYQPLLSNTLAE	457
Qy	431	AKNKLNNETYSR-RSRROAESTTDPMMEM-TGNGAGGEYSSENSITVAQOVAYDNLAIR	488
Db	458	---LYVREYMEQDRKPNATPAPLREAPSANASVERIKTSSIEFARLQFTYNHIORH	513
Qy	489	INNILEDISKAWCRQHRALVWNLKINPTFSVMISYINRNPVSAKRIGDVISNCTIV	548
Db	514	VNDMLGRIAVAWCELQNHLELTWNEARKLNPNNAIASATVGRRVSAARMGLGVMAVSTCPV	573
Qy	549	DOTSVLSHLKRLLSASDEKESRPPVTEKPMNDSTIYKQGVNNEILLTTTYLETQOE	608
Db	574	APDNVIVQNSMR-VSSRPGTCYSRLVSRFVEDQGLLEGOLGENNELRLTRDALEPCTV	632
Qy	609	NTEYFQAKTDMYIKYNEHLKTVPLSSITTLDTFIALNFTLLENVDKVIETLDRSEKR	668
Db	633	GHRRFIFGGGVYEEVAYSHQLSRADVTTVSTFIDLNITWLEDHFVPLEVYTRHEIK	692
Qy	669	LSNVFDIETMFREYNYAQRVGLRKDLDDISTNRNQFVD--AFGLMDDLG-AVGQTV	725
Db	693	DSGLLDYTEVQRRNGLHLRFPADI-DTVIRADANAAMFAGLCAFEPEGMDLGRAVGKVM	751
Qy	726	NAVSGVATLPSIYTGFINFINPFGMLMIIVIGVLFAIYPLTKTKIYETAPIKMIY	785
Db	752	GVVGGV-----SAVSGVSSFSMNPFGALAVGLLVLAVAAFFAFRYVLQQRNPMKALY	807
Qy	786	PEIDKLKEREGKSEIAPISSE-----ELERIVLAMIHQONSHMETKTRKDP	832
Db	808	PLTTKELKTSPPGGVGGEGEAGEGGDEAKLAAREMIRYMALVSAMETERHARKKG	867
Qy	833	KDSIL-TRAQNM-LRKR--SGYSNLKNAE	857
Db	868	TSALLSSKVTNNVLRKRKARYSPLHNE	896
RESULT 23			
AAW00376			
ID	AAW00376	standard; Protein; 904 AA.	
XX			
AC	AAW00376;		
XX			
DT	24-JAN-1997	(first entry)	
XX			
DE	HSV-2 glycoprotein B.		
XX			
KW	HSV-2; glycoprotein B; vaccine; diagnosis; immunoassay.		
XX			
OS	Herpes simplex virus type 2.		
XX			

FH Key Location/Qualifiers
 FT Peptide 1..22
 FT Peptide /label= Sig_peptide
 FT Peptide 18..75
 FT Domain /label= Dominant_type-specific_epitope
 FT Domain 23..744
 FT Domain /label= Extracellular_domain
 FT Domain 745..798
 FT Domain /label= Transmembrane_domain
 FT Domain 799..804
 FT Peptide /label= Cytoplasmic_domain
 FT Peptide 819..904
 FT Peptide /label= Cross-reactive_antigenic_epitope

XX WO9632962-A1.

XX 24-OCT-1996.

XX 19-APR-1996; 96WO-US05316.

XX 21-APR-1995; 95US-0426604.

XX (UYNE-) UNIV NEW MEXICO STATE.

XX Bell R, Goade DE, Jenison S;

XX WPI, 1996-485557/48.

XX New type-specific and cross-reactive Herpes Simplex Virus epitope(s)
 - used for diagnosing HSV-1 and/or HSV-2 infections and in vaccines
 for producing neutralising antibodies

XX Disclosure; Fig 5; 37pp; English.

CC Herpes simplex virus type 2 (HSV-2) glycoprotein B (gB2) (AAW00376)
 and HSV-1 gB1 (AAW00375) include type-specific and cross-reactive
 antigenic epitopes. Recombinant type-specific epitopes can be used
 in immunoassays to distinguish between HSV-1 and HSV-2 infections.
 CC Recombinant cross-reactive epitopes are useful for the diagnosis of
 both HSV-1 and HSV-2 infections. Both types of epitope can be used
 in vaccines.

XX Sequence 904 AA;

Query Match 18.5%; Score 839; DB 17; Length 904;

Best Local Similarity 28.3%; Pred. No. 5.9e-53;

Matches 246; Conservative 164; Mismatches 367; Indels 92; Gaps 27;

Qy 33 TSPNTATWSTESPLTGHYTHDSSHGCRGNNENRDEEQNKNIYGSPTFFYRVCSASG 92

Db 76 TPPPD-----ANATVAAGHATLR--AHLREIKVENADAQ-----FVVCPPPT 115

Qy 93 VGDVFRFOTDHCPCDASD-MVHSEGIILYKQNIIPFMFRVRKYRKVVTSTV----- 144

Db 116 GATVQFEQPRRCPTREPCQNTYEGIAVFKENIAPYKFKATMYKDVTSQVWFGHRS 175

Qy 145 -YNGIYSDSITNOHTFYKIEPWE--TEKMDTIYQCFNSLRNTGNNLLTYVDRDDNMT 201

Db 176 QFMGIFED-----RAPVFEEVIDKINAKGVCRCSTAKYVRNNMETTAFHRDDHETD 226

Qy 202 VFLQPVGDVTPDKVKGQPELYLPGFWGYSRRRTTVNCELMDFARSNPPDFVTA 261

Db 227 MELKPAK-VATRTSRGWHHTDLYKNPSRVEAFHYGTTVNCIVEEDARSVPYDFEFLA 285

Qy 262 TGDTVEMSPFWG-GEDDH-ENKMHKPEFVSVINNY-KVVDYQNRGTVPGLKTRIFLDRE 318

Db 286 TGDFVMSPFYGYREGSHTHTSYAADRFKQVDFYARDLTTKARATSP--TTRNLLTTP 343

Qy 319 EYTLSEKHLKNNMSYCPPLTWKAFYNGIQTEHSGSYHFVANDITASFTT-----SKED 371

Db 344 KFTVAMDVWPKRPVACTMTKQVEDEMLRAEYGGSFSSDAISTTFTTNLTQYSLSRVD 403

Qy 372 MKEFNTTYHCLNEEIKAEIEKKYA-KVNSTHSKYGDLKFKYKTDGGLYLWQPLIQNRLLD 430

```

Db 404 LGD-----CIGRDAREALDRMFARKYNATHIKVGQPOYLLATGFLIAYQPLSLNTLAE 457
Qy 431 AKNKLNETYSR-RSRROAESTTDPMMEM-TGNAGAGEYSSENSITVAQVQYADNLRIR 488
Db 458 ----LYVREYMRQDRKPRNATPAPLRAPASANASVERIKTSSIEFARLOFTYNNHIQRH 513
Qy 489 INNILEDLSKAWCREQHRALVWNLKINPTSVMSMIYNRPVSAGKRGDVISVNCIVV 548
Db 514 VNDMLGRIAVAWCELQNHLETLWNEARKLNPAIASATVGRVRVARMGLGVMAVSTCPV 573
Qy 549 DQTSVSLHLSRLLSASDEKCFSPPTVTFKFMNDSTIYKGLGVNNEILLTITYLETQCE 608
Db 574 APDNVIVQNSMR-VSSRPGTCYSRPLVSFRYEDQGLIEGOLGENNELRLTRDALEPCTV 632
Qy 609 NTEYFQAKTDMYIKNYEHLKTVPLSSITLDTFTIALNFTLLENVDFKVIETYRDEKR 668
Db 633 GHRRYFIFGGYVYFEYAYSHQLSRADVTTVSTFIDLNTMLEDFEFLVPLEVYTRHEIK 692
Qy 669 LSNVFDIETMFREYNYAQRVSGRLKDLLSLTRNQFVD--AFGSLMDLGL-AVGQTVV 725
Db 693 DSGLLDYTEVQRNQLHDLRFADI-DTVIRADANAAMPAGLCALFEFGGDDIGRAVGKVM 751
Qy 726 NAVSGVATLFSISITVTFINFKNPPFGGMLMIIVVIGVLFALFYFLTKTKIYETAPIKMIY 785
Db 752 GVVGGVV----SAVSGVSSFMSPFGALAVGLLVLAGLVAAFPAPRYVLQLRNPKALY 807
Qy 786 PEIDKLEREGKSEIAPISSE-----ELERIVLAMIHQONSHMETKTRKOP 832
Db 808 PLTTKELKTSDPGGVGGEAGEGGGFDEAKLAAREMIRYALVSAMERTEHKARKKG 867
Qy 833 KDSIL-TQAQNM-LRKR--SGYSNLKNAE 857
Db 868 TSALLSSKVTNWLKRNKARYSPLNED 896

RESULT 24
AAW72193
ID AAW72193 standard; Protein; 904 AA.
XX
AC AAW72193;
XX
DT 13-JAN-1999 (first entry)
XX
DE HSV-2 strain SB5 Contig ID 15 ORF#28 protein.
XX
KW HSV-2 strain SB5; immunological response induction; therapy;
KW antiviral identification; viral protein inhibitor.
XX
OS Herpes simplex virus type 2.
XX
PN W09820016-A1.
XX
PD 14-MAY-1998.
XX
PF 31-OCT-1997; 97MO-US20016.
XX
PR 09-JUN-1997; 97US-0049018.
PR 04-NOV-1996; 96US-0030279.
XX
PA (SMIK ) SMITHKLINE BEECHAM CORP.
XX
PI Chan JY, Dabrowski-Amaral CE, Delvecchio AM, Dillon SB;
PI Esser KM, Leary JJ;
XX
DR WPI; 1998-286847/25.
DR N-PSDB; AAV62176.
XX
PT Herpes simplex virus type-2 sequences - useful in, e.g. prevention
PT and treatment of infection or inducing immunological response in
PT mammal
XX
PS Claim 10; Page 121; 748pp; English.

```

```

XX
CC This sequence represents a Herpes simplex virus type-2 (HSV-2) protein
CC sequence of the invention. This sequence was isolated from a HSV-2 strain
CC SB5 (deposited as ATCC VR-2546) DNA fragment designated Contig ID 15.
CC Based on homology, this sequence is a glycoprotein B precursor.
CC The proteins can be used for the treatment or prevention of disease, to
CC induce an immunological response in a mammal or to identify inhibitors,
CC activators or novel antivirals. Antagonists of the proteins can be used
CC to inhibit a viral polypeptide. The DNA sequence or a vector containing
CC it can also be used to induce an immunological response in a mammal.
XX
SQ Sequence 904 AA;
Query Match 18.5%; Score 839; DB 19; Length 904;
Best Local Similarity 28.3%; Pred. No. 5.9e-53;
Matches 246; Conservative 164; Mismatches 367; Indels 92; Gaps 27;
Qy 33 TSPNTATWSTESPLTGHYTHDSHGGERGNENRDSSEQNKNIYGSPTFPYRVCSASG 92
Db 76 TTPPD-----ANAIVAAGHATLR--AHLREIKVENADAQ-----FVCCPPT 115
Qy 93 VGVFRFQTDHVCDDASD-MVHSEGILLIYKQNIIPMFVRKYRKVVTTSTV----- 144
Db 116 GATVWQPEQRRCPTRPEQONYTEGIAVPEKENIAPYKFKATMYKDKVTVSQVWFGHRS 175
Qy 145 -YNGIYSDSITNQHTFYKSTIEPWE--TEKMDTIYQCNSRLNTGGNLLTYVDRDDINMT 201
Db 176 QFMGIFED-----RAPVPEEVIDKINAKGVCSTAKYVRNNMETTAFHRDDHETD 226
Qy 202 VFQPPQVGVTPDKVYKSGQPELYLEPCWFGVSGYRRRTTNCVCELMDFARSNPPDFPVTA 261
Db 227 MELKPAK-VATRTSRGHTTDLKYNFSRVEAFHYGTGTVNCIVEEDARSVYDFEFLA 285
Qy 262 TGDVTVMSPFWS--GEDDH-ENKMHKXPWFVSVNINNY-KVVDYQNRGTVPGLKTRIFLDRE 318
Db 286 TGDVFMSPFYGVRGSGHTEHTSYAADRFKQVDFYARDLTTKARATSP--TTRNLLTTP 343
Qy 319 EYTLSEKHLKMSYCPGLTLKAFYNGIQTEHSGSYHFVANDITASPTT-----SKED 371
Db 344 KFTVAVDWPKRPVACTMTKQVDEMLRAEYGGSPFRSSDAISTTFTTNLTQVLSRVD 403
Qy 372 MKEFTNTYHCLNEEIKAEIEKYYA-KVNSTHSGYGLKYPKFTDGGLYLVWQPLQIRLLD 430
Db 404 LGD-----CIGRDAREALDRMFARKYNATHIKVGQPOYLLATGFLIAYQPLSLNTLAE 457
Qy 431 AKNKLNETYSR-RSRROAESTTDPMMEM-TGNAGAGEYSSENSITVAQVQYADNLRIR 488
Db 458 ----LYVREYMRQDRKPRNATPAPLRAPASANASVERIKTSSIEFARLOFTYNNHIQRH 513
Qy 489 INNILEDLSKAWCREQHRALVWNLKINPTSVMSMIYNRPVSAGKRGDVISVNCIVV 548
Db 514 VNDMLGRIAVAWCELQNHLETLWNEARKLNPAIASATVGRVRVARMGLGVMAVSTCPV 573
Qy 549 DQTSVSLHLSRLLSASDEKCFSPPTVTFKFMNDSTIYKGLGVNNEILLTITYLETQCE 608
Db 574 APDNVIVQNSMR-VSSRPGTCYSRPLVSFRYEDQGLIEGOLGENNELRLTRDALEPCTV 632
Qy 609 NTEYFQAKTDMYIKNYEHLKTVPLSSITLDTFTIALNFTLLENVDFKVIETYRDEKR 668
Db 633 GHRRYFIFGGYVYFEYAYSHQLSRADVTTVSTFIDLNTMLEDFEFLVPLEVYTRHEIK 692
Qy 669 LSNVFDIETMFREYNYAQRVSGRLKDLLSLTRNQFVD--AFGSLMDLGL-AVGQTVV 725
Db 693 DSGLLDYTEVQRNQLHDLRFADI-DTVIRADANAAMPAGLCALFEFGGDDIGRAVGKVM 751
Qy 726 NAVSGVATLFSISITVTFINFKNPPFGGMLMIIVVIGVLFALFYFLTKTKIYETAPIKMIY 785
Db 752 GVVGGVV----SAVSGVSSFMSPFGALAVGLLVLAGLVAAFPAPRYVLQLRNPKALY 807
Qy 786 PEIDKLEREGKSEIAPISSE-----ELERIVLAMIHQONSHMETKTRKOP 832
Db 808 PLTTKELKTSDPGGVGGEAGEGGGFDEAKLAAREMIRYALVSAMERTEHKARKKG 867

```

QY 833 KDSIL-TRAQNM-LRKR--SGYSNLKNAE 857
 Db 868 TSALLSSKVTNNVLKRKARYSPLHNE 896

RESULT 25
 AAB74442
 ID AAB74442 standard; protein; 904 AA.
 AC AAB74442;
 XX 29-MAY-2001 (first entry)
 DT 29-MAY-2001 (first entry)
 DE Herpes simplex virus 2 glycoprotein B.
 KW HSV-1; HSV-2; glycoprotein B; gb; transmembrane envelope glycoprotein;
 KW antigenic epitope; diagnosis; vaccine.
 XX Herpes simplex virus type 2.
 OS Herpes simplex virus type 2.
 XX US6197497-B1.
 PN 06-MAR-2001.
 PD 19-APR-1996; 96US-0632537.
 PF 21-APR-1995; 95US-0426604.
 PR (UYNE-) UNIV NEW MEXICO STATE.
 PA Goade DE, Bell R, Jenison S;
 PI WPI; 2001-256360/26.
 DR Continuous, isolated, antigenic polypeptide segment of herpes simplex
 PT virus (HSV) glycoprotein B1 or B2, useful in serodiagnostic
 PT immunoassays for distinguishing HSV-1 infection from HSV-2 in a human
 XX Claim 1; Fig 5; 23pp; English.
 CC The present invention provides antigenic peptides from herpes simplex
 CC virus type 1 (HSV1) and 2 (HSV2) glycoprotein B (gb) which can be used in
 CC the diagnosis of HSV infection, and identification of subtype, and in
 CC vaccines to protect against HSV. The present sequence is the HSV-2 gb
 CC protein.
 CC Note: The present sequence is mentioned in claim 8 as being the HSV
 CC glycoprotein B2.
 SQ Sequence 904 AA;

Query Match
 Best Local Similarity 18.5%; Score 839; DB 22; Length 904;
 Matches 246; Conservative 164; Mismatches 367; Indels 92; Gaps 27;

QY 33 TSPNTATWSTESPLTGHYTHDSHGERNNENRDSBEQNKNIYSPSTFPYRVCASG 92
 Db 76 TTPPD-ANATVAAGHATLR-AHLREIKVENADAQ-----FYVCPPT 115
 QY 93 VGDVFRQTDHVCPSAD-MVHSEGLIYKQNIIPFMRVKRYKVVTTSTV----- 144
 Db 116 GATVVQFEQPRCPTREPCQNTGEGIAVVFKEINAPIYKFKATMYKYKDVTSQVWFGHRY 175
 QY 145 -VNGIYSDSITNQHFTYKSIPEW--TEKMDIYOCFNSLRNTGNNLLTYVDRDDINMT- 201
 Db 176 QMGIFED-----RAPVPFEVIDKINAKGVCSTAKYVRNNMETTAFHRDDHETD 226
 QY 202 VFLQPVGDVTPDKRYGSOPELYLEPGFWGYSRRRTTNCCLMDFMARSNPPEFFVTA 261
 Db 227 MELKPAK-VATPSTSGMHTDLKYNPSRVEAFHRYGTTCVCEVDARSYPDFEFVLA 285
 QY 262 TGDVTVMSPFWS-GBDDH-ENKMHKPKFWFSVINNY-KVDYQNRGTVPGLKTRIFLDRE 318
 Db 286 TGDVYMSFPFYGYREGSHTENTSAAADRFPKQVDFYARDLTTKARATSP--TTRNLLTTP 343

319 EYTLSEKHLKMSYCPPLTLWKAFFNGIOTBHSYHFEVANDITASFTT-----SKED 371
 Db 344 KFTVANDWVPRPAVCTMTKQOEVDMLRABYGSGSFRESSDAISTFTFTNTLTQVLSLSD 403
 QY 372 MKEFTNTYHCLNEEIKABIEKKYA-KVNSTHSKYGLDKYFKTGDGLYLVWQPLQNRLLD 430
 Db 404 LGD-----CIGRDAREAIDRMFARKYNATHIKVGQPOYYLATGGFLIAYQPLLSNTLAE 457
 QY 431 AKNNLNNETYSR-RSRRQAEESTTDPMMEM-TCNGAGGEYSSENSITVAQVAYDNLIR 488
 Db 458 ----LYVREYMRQDRKPRNATPAPLRAPASANASVERIKTTSIEFARLOFTYNNHIORH 513
 QY 489 INNILEDLSKAWCREQHRALVWNELSKINPYSVMYINRPVSAKRIGDVISVNCIWW 548
 Db 514 VNDMLGRIAVAWCLQNHLELTWNEARKLNPNATASATVGRVSRMLGDVWAVSTCVPV 573
 QY 549 DOTSVLHKSURLLSASDEKCFSPPTVTFKPMNDSTIYKGLGVNNEILLTTTLYLETCOE 608
 Db 574 APDNVIVQNSMR-VSSRPGTCYSRPLVSFRYEDQGLIEGQLEGGNNELRLTRDALEPCTV 632
 QY 609 NTEYFQAKTDMYIKNYEHLKTVPLSSIITLDTFIALNFTLLENVDFKVIETYRDEKR 668
 Db 633 GHRRYFIFGGGVYFEEYAYSHQLSRADVTTVTFIDLNITMLDEHFPVLEVYTRHEIK 692
 QY 669 LSNVFDIETMFREYNYAQRVSLRKLDDLSTNRNQFVD--AFGSLMDDLGL-AVGQTVV 725
 Db 693 DSGLLDYTEVORRNQLHDLRFADI-DTVIRADANAAMFAGLCAFEFGMDLGRVGVKVM 751
 QY 726 NAVSGVATLFSSIVTGFINFINKPFGGMLMIIVVIGVLFAIYFLTKTKIYETAPIKMIY 785
 Db 752 GVVGGVV----SAVSGVSSFMSPFPGALAVGLLVLAGLVAAPFAFRYVVLQQRNPMKALY 807
 QY 786 PEIDKLKEREKSEIAPISEE-----ELERIVLAMHIHQNSHMETKTRKDP 832
 Db 808 PLTTKELKTSPPGGVGGEGEGGEGGEGGEGGEGGEGGEGGEGGEGGEGGEGGEGGEGG 867
 QY 833 KDSIL-TRAQNM-LRKR--SGYSNLKNAE 857
 Db 868 TSALLSSKVTNNVLKRKARYSPLHNE 896

RESULT 26
 AAP60244
 ID AAP60244 standard; protein; 903 AA.
 XX AAP60244;
 AC AAP60244;
 XX 25-MAR-2003 (updated)
 DT 31-JUL-1991 (first entry)
 DE Herpes simplex virus glycoprotein gb.
 KW HSV; gb glycoprotein; vaccine.
 OS Herpes simplex virus.
 XX EP170169-A.
 XX 05-FEB-1986.
 XX 19-JUL-1985; 85EP-0109042.
 XX 20-JUL-1984; 84JP-0151766.
 PR 11-DEC-1984; 84JP-0262465.
 XX (KAGA) CHERO SERO THERAPEUTIC RES INST.
 XX Noraki C, Makizumi K, Kino Y, Eto T, Ohtomo N;
 WPI; 1986-036935/06.
 DR N-PSDB; AAN60195.
 XX


```

Db 170 --RAPVPE- EYMDKINAGVCKSTAKYVRNNWESTAFHRDDHESDMALKPAKAAT-RTS 225
QY 216 RYGSQPELYLEPGWFWGVRRTTVCNCELMDFARNPPDFVATGDTVMESFFWGE 275
Db 226 RGMHTDLYNPARVEAFHRYGTTVCNCEIVEARSVPYDFVLATGDFVYMSPFYGR 285
QY 276 D-DH-ENKMKHKEPWFVSVINNY-KVVDYQNRGTVPGLGTRIPLDREEVYTLSEKHLKMS 332
Db 286 DSGHGEHTAYADRFRQVDGYERDLSTGRRAAPV--TRNLLTPKFTVGHWDWAPKPS 343
QY 333 YCPLTLKAFYNGIQTEHSGSYHFVANDITASFTT-----SKEDMKEFNNTYHCLNEE 385
Db 344 VCTLTKWREDEWMLRAEYSPFRSSAALSTFTTANRTEYALSRLDLAD-----CVGRE 397
QY 386 IKAIEKKY-AKVNTHSKYGLKFKTDGGLYLVWQPLIQNRLLDAKKNLNEYYSRS 444
Db 398 AREAADRFIRRYNGTHVKVGQVYLYLATGGFLIAYQPLLSNALV-----ELYVREL 449
QY 445 RRAESTTDPMMEMTGNAGGE-----YSSENSITVAQVOYAYDNL 486
Db 450 VR-----EQTRPAGGPGGAATPGSPDPPSVVERIKTSSVFARLQFYDHIQ 499
QY 487 TRINNLEDLSKAWCREQHRALVWNELSKINPTSMGSMYNNRPSAKRIGDVISNCCI 546
Db 500 RHVNDMLGRIATAWCELQNRLLTWNERRLNPGAISATVCGRRYSARMGLGDMVASTCV 559
QY 547 VVDQTSVSLKSLRLLSASDEKCFRPPVTFKPMNDSTIYKQLGVNNEILLTYYLET 606
Db 560 VPAPDNVIMQNSMR-VAARPQTCYSRPLVSFRYADGPLVEQLGEDNEIRLERDALP 618
QY 607 QENTYVFOAKDMYIKNYEHLKTVPLSSITTLDTETALNFTLENVDFKVIELYTRDE 666
Db 619 TVGHRRYTFGAGYVFEYAYSHQGLRADVTIVTFINLNTMLDEHFEVPLEYTRQE 678
QY 667 KLSNVFDIETFREYNYAQRVSLGRKDLDDLSTNRNQVDAFGLSMDLDDGAVGQTVN 726
Db 679 IKDSGLLDVTEVORRNQLHALFADI-DTVIKADAAHALFAGLY-SFPEGLGDVGRAVK 736
QY 727 AVSGVATLFSSIVTGFINKPFGGMLMIIVVIGVLFAYIPLTKTKIYETAPKMIYP 786
Db 737 VMGIVGVVSAVSGVSSFLSPFGALAVGLLVLAGLAAAFPAFRYVMRLQRPNMKALYP 796
QY 787 EIDKLKREGKGEIAPIS-----EEELERIVLA-----MHTHQNSHMETKTRDP 832
Db 797 ----UTTKELSDGAPLAGGGEDGAEEDFDEAKLAQAREMIRYMALVSAMERTEHKARKG 852
QY 833 KDSIL---TRAQNMLRKRSYSLNKNAESVEM 861
Db 853 TSALLSAKVTDAVMEKRAPRYSPLRDTDEEL 885

RESULT 28
AAW34552
ID AAW34552 standard; Protein; 904 AA.
XX
AC AAW34552;
XX
DT 26-FEB-1998 (first entry)
XX
DE Herpes simplex virus type 1 glycoprotein B.
XX
KW Herpes simplex virus; HSV; type I; type II; glycoprotein B; vaccine;
KW membrane protein; virus-specific glycoprotein;
KW transmembrane anchor region.
XX
OS Herpes simplex virus type 1.
XX
FH Key Location/Qualifiers
FT Domain 726..795
FT Region /label= membrane spanning_region
FT /note= "putative"

```

```

XX US5648079-A.
PN 15-JUL-1997.
PD 08-DEC-1994; 94US-0351875.
PF 20-OCT-1986; 86US-0921730.
XX 06-APR-1984; 84US-0597784.
PR 20-SEP-1990; 90US-0587179.
PR 21-DEC-1992; 92US-0993415.
PR 18-OCT-1993; 93US-0138717.
XX (CHIR ) CHIRON CORP.
XX Burke RL, Pacht C, Valenzuela PDT;
PI WPI; 1997-372022/34.
XX N-PSDB; AAT93650.
XX Vaccines against herpes simplex virus infection - containing
PT recombinant HSV glycoprotein B
XX
XX Disclosure: Fig 4A-F; 33pp; English.
XX The present sequence is from the Herpes simplex virus (HSV) type I
CC strain Patton and represents a glycoprotein B. HSV is a double stranded
CC virus packaged within an icosahedral nucleocapsid enveloped within a
CC membrane. The membrane includes a number of virus-specific
CC glycoproteins, with glycoprotein B being one of the most abundant.
CC Glycoprotein B from both HSV type I and type II are cross reactive. New
CC vaccines comprising recombinantly produced glycosylated glycoprotein B
CC that has a deletion of at least part of the transmembrane anchor region,
CC in combination with a carrier and an adjuvant have been produced. The
CC vaccines are used for immunising humans against HSV (HSV type 1 or
CC HSV type 2) by vaccination before or after primary infection with HSV.
XX
SQ Sequence 904 AA;
Query Match 18.4%; Score 838; DB 18; Length 904;
Best Local Similarity 28.0%; Pred. No. 7e-53;
Matches 243; Conservative 166; Mismatches 368; Indels 92; Gaps 26;
QY 33 TSPNTATWSTESPLTGHYTHDSSHGRCNNENRDSSEQNKNIYGSPTFPYRVCASG 92
Db 76 TTPPD-----ANATVAAGHATLR--AHLREIKVENADAQ-----FYVCPPT 115
QY 93 VGDVRFOTDHVCPDASD-MVHSEGLLIYKQNIIPPMFRVRKYRKVVTSTV----- 144
Db 116 GATVQPEQPRCPTRPQNYTEGIAVVKENIAPYKFKATMYKDVTVSQVWFGHYS 175
QY 145 -YNGIYSDSITNQHTFYKSIPEWE--TERKMDTIYQCFNSLRNLNTGGNLLTYVDRDDINMT 201
Db 176 QFMGIFED-----RAPVPEEVIDKINAKGVCSTAKYVRNNMETTAFHRDDHETD 226
QY 202 VFLOPVDGVTDPVKYGSQPELYLEPGWFWGYSYRRTTVCNCELMDFARNPPDFVTA 261
Db 227 MELKPAK-VATRTSGRWHTTDLKYNPSRVEAFHRYGTTVCNCEIVEARSVYPYDFVLA 285
QY 262 TGDVEMSPFWS-CEDDH-ENKMKHKEPWFVSVINNY-KVVDYQNRGTVPGLKTRIFLDRE 318
Db 286 TGDVYMSPFYGYREGSHTSHTSVAADRFXQVDFYARDLTTKARATSP--TTRNLLTTP 343
QY 319 EYTLSEKHLKMSYCPPLTLKAFYNGIQTEHSGSYHFVANDITASFTT-----SKED 371
Db 344 KFTVADWVPRPAVCTMTKQVEDEMLRAEYSGFRSSDAISTTFTTNTLTQYSLSRVD 403
QY 372 MKEFNTHYHCLNEEIKAEIEKKYA-KVNSTHSGYGLKFKTDGGLYLVWQPLIQNRLLD 430
Db 404 LGD-----CIGRDAREAIDRMFARKYNATHIKVGQPYLYLATGGFLIAYQPLLSNTUAE 457
QY 431 AKNKLNNETYSR-RSRQAESTTDPMMEM-TGNCAGGGEYSSENSITVAQVOYAYDNLRI 488

```

Db 458 ----LYVREYREODRPNATPAPLREAPSANASVERIKTSSIEFARLQFTYNNHQH 513
 Qy 489 INNILEDLSKAWCREQRAALVWNELSKINTPSVMSIYNRPVSAKRIGDVISYSCIVV 548
 Db 514 VNDMLGRIAVAWCBLQNHLETLWNEARKLNPAIASATVGRVSARMGDDVMAVSTCPV 573
 Qy 549 DOTSVLHKSURLLSASDEKCFSPRPVTFKFMNDSTIYKGLGVNNEILLTTTYLETQ 608
 Db 574 APDNVIVQNSMR-VSSRPGTCSYRPLVSFRYEDQGLIEGQGENNELRLTRDALEPCTV 632
 Qy 609 NTEYFFQAKTDMYIKYVEHLKTVPLSGSITTLDTFIALNFTLLENVDFKVIETROEKR 668
 Db 633 GHRRYFIYFGGYYVEEYAYSHQSLRADVTTVSTFIOLNITMLDHEBFVPLEVTRHEIK 692
 Qy 669 LSNVFDIETPREYNYAQRVSGRLKDLDLSTNRNQFVD--AFGSLMDDLQ-AVGQTVV 725
 Db 693 DSGLLDTEVORRNQLHDLRFADI-DTIVIRADANAAMFAGLCAFEFGMGDDLGRAVGKVM 751
 Qy 726 NAVSGVATLFSIVTGFNFINKPFGGMMLIIVVIGVLFAIYFLTKTKIYETAPIKMIY 785
 Db 752 GWVGWV----SAVSGVSSFMNPFGLAVGLLVLAGLVAAFFAFRYVLQLRNPMKALY 807
 Qy 786 PEIDKLKEREKSEIAPISEE-----ELERIVLAMIHOONSHMETKTRKDP 832
 Db 808 PLTTKELKTSDPGGVGGEAGEGAGGGFDEAKLAAREMIRYMALVSAMERTEHKARKKG 867
 Qy 833 KDSIL-TRAQNMW---RKRSGYSLKNAE 857
 Db 868 TSALLSSKVTNVLAKNKARYSPLHNE 896

RESULT 29

AA711136
 ID AAP711136 standard; Protein; 907 AA.

AC AAP711136;

CC 25-MAR-2003 (updated)

DT 30-APR-1991 (first entry)

XX Herpes Simplex Virus-2 gB from p52BXX.

KW Vaccine; prophylaxis; treatment; Herpes Simplex Virus-1;
 glycoprotein; gB.

OS Herpes simplex virus type 2 HG52.

XX US4642333-A.

XX 10-FEB-1987.

XX 20-JUN-1984; 84US-0622496.

XX 20-JUN-1984; 84US-0622496.

XX (PERS/) PERSON S.

XX Person S;

XX WPI; 1987-056354/08.

XX N-PSDB; AAN71399.

XX Amino acid chain of glyco:protein B of HSV-1 and 2 - prepd. as
 PT recombinant and used for vaccines for herpes simplex virus types 1
 PT and 2.

XX Example; Table 2; 16pp; English.

XX

CC A pure non-glycosylated amino acid (AA) chain comprising a sequence
 CC corresponding to that occurring in glycoprotein B of HSV-1 or HSV-2
 CC virus which is antigenic to HSV-1 of HSV-2, which contains no more
 CC than 750 AA residues, and which includes AA residues 135-649
 CC inclusive is claimed. It can be used to produce vaccines for

CC prophylaxis and treatment of HSV-1 and HSV-2.
 CC See AAN71303 for the HSV-1 sequence.
 CC (Updated on 25-MAR-2003 to correct PR field.)

XX Sequence 907 AA;

Qy Query Match 18.4%; Score 838; DB 8; Length 907;

CC Best Local Similarity 28.8%; Pred. No. 7e-53;

CC Matches 241; Conservative 163; Mismatches 347; Indels 86; Gaps 26;

Qy 65 ENRDSBQNKNIYGSPTFFYRVCSAGVGDFRQTDHVCPPDASD-MVHSEGLLIYKQ 123

Db 105 ENADAQ-----FYVCPPTGATVVOFEQPRECPTRPEQONTTEGIAVVFKE 150

Qy 124 NIIPFPRVRKYKVVTTSV-----YNGIYSDSITNQHTFYKSEIWE--TEKMDT 173

Db 151 NIAPYKFKATMYKDVTVSQVWFGRYSQFMGIFED-----RAPVPEEYIDKINA 201

Qy 174 IYOCFNSLRLLNTGNNLLTYVDRODINNVTFLQPVGVTPDVKRYGSOPELYLEPGWFGS 233

Db 202 KGYCRSTAKVVRNMETAFHRDDHETDMELKPAK-VATRTSGWHTTDLKYNPSRVEAF 260

Qy 234 YRRRTTVNCELMDFARSNPFPFVYATGDTVMESPFW-S-GEDDH-ENKNHEKPFVSV 291

Db 261 HRYGTTVNCIVEVDARSVVPYDEFVLATGDFVMSPPFYCYREGSHEHTTYAADRQKV 320

Qy 292 INNY-KVVDYQNRCTVPLGKTRIFLDREYITLSWEKHLKNMSYCPLTLWKAFYNGIOTEH 350

Db 321 DGFYARDLTTKARATAP--TTRNVLTTPKFTVAMDVPRKPAVCTMTKQEVDEMLRAEY 378

Qy 351 SGSYHFVANDITASFTT-----SKEDMKEFNTYHCLNEEIKABIEKKYA-KVNSTHS 402

Db 379 GGSRESSDAISTFTTNLTYSLSRVDLGD-----CIGRDAREALDRMPARKYNATHI 432

Qy 403 KYGDLKYFTDGLYLWQPLIQNRLLDKAKLNNETYSR-RSRQAEESTTDDMEMEM-TG 460

Db 433 KVGOPYQATGGFLIAYQPLLSNTLAE---LYVREYMRQESKRKPGNATPAPLREAPS 488

Qy 461 NGAGGEYSSENSITVAQVQYADNLRIRINNILEDLSKAWCREQRAALVWNELSKINPT 520

Db 489 NASVERIKTSSIEFARLQFTYNNHQHVNMDLGRIVAVMCELOHNLTLWNEARKLNPN 548

Qy 521 SVMSMIYNRPVSAKRIGDVISYSCIVVDOTSLSLHLSRLLSASDEKCFSPRPVTFKPM 580

Db 549 ATASATVGRVSARMGDDVMAVATCPVAPDNVIVQNSMR-VSSRPGTCSYRPLVSFRYE 607

Qy 581 NSTIYKGLGVNNEILLTTTYLETQENTYFFQAKTDMYIKYVEHLKTVPLSSITTL 640

Db 608 DQGPLIEGQGENNELRLTRDALEPCTVGHRRYFIYFGGYYVVEYAYSHQSLRADVTTV 667

Qy 641 DTFIALNFTLLENVDFKVIETROEKRLSNVDPDIETMFREYNYAQRVSGRLKDLDDL 700

Db 668 STFDILNITMLDHEFVPLEVTRHEIKDSGLLDYTEVQRRNQLHDLRFADI-DTIVIRAD 726

Qy 701 TNRNQFVD--AFGSLMDDLQ-AVGQTVVNAVSVATLFSISIVTGFNFINKPFGGMMLII 757

Db 727 ANAAMFAGLCAFEFGMGDDLGRAVGKVMGVGVV----SAVSGVSSFMNPFGLAVGL 782

Qy 758 VVIGVLFAYFLTKTKIYETAPIKMIYPEIDK-LK-----EREKSEIAPISEE 807

Db 783 LVLAGLVAAFFAFRYVLQLRNPMKALYPLTTKELKLPDPPGGVGGEAGEGAGGGFDEAK 842

Qy 808 L---ERIVLAMIHOONSHMETKTRKDPKDSIL-TRAQNMW---RKRSGYSLKNAE 857

Db 843 LAAREMIRYMALVSAMERTEHKARKKGTSSALLSSKVTNVMVRKNRKARYSPLHNE 899

RESULT 30

AA411779

ID AAR41779 standard; Protein; 904 AA.

XX AAR41779;

XX


```
Db 116 GATVQFEQPRRCPTREQNTYEGIAVVFVKENIAPYKFKATMYKDVTSQVWFGRHYS 175
Qy 145 -YNGIYSDSITNQHTFYKISPEWE--TEKMDTIYOCFNSRLNTGNNLLTYVDRODINWT 201
Db 176 QWGIPEF-----RAPVFEEVIDKINAKGVCRTAKVVRNNMETTAFHRDHD 226
Qy 202 VFLQPVGVTPDVKRYGSOPELYLEFGWFWGYSRRRTTNCBMDMFARSNPPFFDFTVA 261
Db 227 MELKPAK-VATRTSRGWHTTDLKPNPSRVEAFHRYGTTVNCIVEEVDARSVYPDEFVLA 285
Qy 262 TGDVTMSFPWS-GEEDH-ENKHEKPFVSVINNY-KVDYQNRGTVPGLKTRIFLDRE 318
Db 286 TGDVFMSPFYGREGSHTSHTSYAADREFQKQDGFYARDLTTKARATSP--TTRNLLTTP 343
Qy 319 EYTLSEKHLKNNMSPCLTLWKAFFNGIOTHSYGVHVFANDITASFTT-----SKED 371
Db 344 KTVAVDWPKPAVCTMTKQVEIDILRAEYGGFRSSDAISTFTTNLTQYLSRVD 403
Qy 372 MKEFNNTYHCLNBEIKAEIKKYA-KVNSTHSGYGLKYFKTDGGLYLWQPLIQNRLLD 430
Db 404 LGD-----CIGRDAREADRMFAKYNATHINVGQPYLATGGFLIAYQPLLSNTLAE 457
Qy 431 AKNKLNNETYSR-RSRQRESTDPMEM-TGNGAGGEYSSENSITVAQVQVAYONLIR 488
Db 458 ----LYVREYMEQDRKPNRATPAPLREAPSANASVERIKTTSIEFFARLQFTYNIQRH 513
Qy 489 INNILEDLSKACREOHRAALVNNELSKINPTSVMSIYNNRVPVSARKIGDVISVNCIVV 548
Db 514 VNDMLGRIAVACELQNHLETLWNEARKUNPNAIASATVGRRRVARMUGDVNAVSTCPV 573
Qy 549 DQTSVSLHSLRLSGLSASDEKFSRPPVTPKFMNDSTIYKQGLGVNNEILLTYYLETQOE 608
Db 574 APDNVIVQNSMR-VSSRPCTCYSRPLVFRYEDQGLIEGQLGNNELRLTRDALEPCTV 632
Qy 609 NTEYFPAQTDYIYKNYEHKTVPLUSSITLDTFIALNFTLLENVDKVIELYTRDEKR 668
Db 633 GHRRYFIFGGVYFVEYAYSHQSRADVTTVSTFDLNTMLDEHFEVPLESVYTRHEIK 692
Qy 669 LSNVFDIETMFREYNYAORVSLRKDLDDLSTNRNQFVD--AFGLMDDDLQ-AVGQTVV 725
Db 693 DGLLDYTVQRNQLHDLRFADI-DTVIRADANAMFAGLCAFFEGMGDLGRAVGKVM 751
Qy 726 NAVSGVATLFSISITGFINFKPNPGMLMIIVIGVLPAIFYLTKTKTIYETAPIKMIY 785
Db 752 GVVGWV---SAVSGVSFMSNPPGALAVGLLVLAGLVAFAFFRYVLQLQRNPMKALY 807
Qy 786 PEIDKLEREGKSEIAPISEE-----ELERIVLAWHIHQONSHMETTKRDP 832
Db 808 PLTTKELKTSDPGGVGGEAGEAGGGFDEAKLAAREMIRYMALVSAAMERTHKARKKG 867
Qy 833 KDSIL-TRAQNM-LRKR--SGYSNLKNAE 857
Db 868 TSALLSKVTNMLVRKNRKARYSPLNED 896

RESULT 33
AAR22615
ID AAR22615 standard; Protein; 868 AA.
XX
AC AAR22615;
XX
DT 25-MAR-2003 (updated)
DT 21-OCT-1992 (first entry)
XX
XX Varicella-zoster virus polypeptide.
XX
XX VZV; protective; immunogenic; gB glycoproteins; chickenpox;
XX
XX Varicella zoster virus.
XX
XX EP482671-A.
XX
```

```
PD 29-APR-1992.
XX
XX 31-JUL-1986; 91EP-0120701.
XX
PR 02-AUG-1985; 85US-0762001.
XX
XX (MERI ) MERCK & CO INC.
XX
XX Davison AJ, Ellis RW, Keller PM, Lowe RS;
XX
XX WPI; 1992-142773/18.
XX
XX N-PSDB; AAQ24233.
XX
XX Vaccine against varicella-zoster virus (VZV) - produced from gene
XX encoding outer surface viral protein
XX
XX Claim 1; Page 9; 18pp; English.
XX
XX Chickenpox is caused by varicella-zoster virus (VZV), a member of
XX the herpesvirus gp. VZV has five major glycoproteins on its surface
XX which are the prods. of three different genes, gA, gB and gC.
XX Monoclonal antibodies to gA and gB display complement-independent
XX neutralisation and the monoclonal antibodies of gC display
XX complement-dependent neutralisation. Cytoplasmic RNAs were prepd.
XX from VZV-infected MRC-5 cells. The RNAs encoded by the different
XX VZV HindIII fragments were selected by hybridisation to cloned VZV
XX HindIII DNA fragments (J.R.Ecker & R.W. Hyman, Proc. Natl. Acad.
XX Sci. USA 79:156 (1982)) bound to nitrocellulose. These RNAs were
XX translated in a rabbit reticulocyte lysate. The polypeptide prods.
XX were immunopptd. by polyclonal monospecific guinea pig antibodies
XX raised to gB purified by monoclonal antibody affinity
XX chromatography. By this analysis it was found that a 100 kD in
XX vitro translation prod. from mRNA selected by the VZV-HindIII-D
XX fragment could be immunopptd. by the anti-gB antibodies which
XX neutralise viral infectivity. The polypeptide may be used to
XX react with human convalescent zoster sera and with monospecific
XX antisera which neutralise viral infectivity. The polypeptide or
XX fragments of it are useful for the prepn. of vaccines or diagnosis
XX of VZV infection.
XX (Updated on 25-MAR-2003 to correct PF field.)
XX
XX SQ Sequence 868 AA;
XX
XX Query Match 18.3%; Score 832; DB 13; Length 868;
XX Best Local Similarity 26.7%; Pred. No. 1.8e-52;
XX Matches 236; Conservative 168; Mismatches 365; Indels 116; Gaps 27;
XX
Qy 40 TWSTESPLTGHYTHDSHGERGNENRDSSEQNKNIYSPSTPFYRVCSAGYGVDPFRF 99
Db 23 TQSEDIITRSHLGDG-----EIREAIHKQDAETKPT---FYVCPPTGTVRL 70
Qy 100 QTDHVCDDASDVH-----SEGILLIYQNIIIPMFVRKYRKVTVTSTVNGIYSDSIT 154
Db 71 EPTRTCPD-----YHLGKNFTGIAVYKFNIAAAYKFKATVYKDVIVSTAWAGSSYTOIT 126
Qy 155 NOHTFYKISPEWE--TEKMDTIYOCFNSRLNTGNNLLTYV-----DRDINNVT 202
Db 127 NRYADRVPIPVSEITDTIDKFGK-----SSKATYVNNHKVAFNEDKKNQDMPL 177
Qy 203 FLQPVGVTPDVKRYGSOPELYLEFGWFWGYSRRRTTNCBMDMFARSNPPFFDFTVA 262
Db 178 IASKYNSV--GSKAWHTTNDTYWVAG--TPGYRTGTSGVNCIIIEVEARSIFPDSFGLST 234
Qy 263 GDTVEMSPFWGSD-----DHEN-----KXHEKPFVSVINNYKVDYQNRGTVPGLKTRIF 314
Db 235 GDIIYMSPPFLGRDGAHREHSNYAMDRFHQ-----FEGYRQRDLDTLALLE--PAARNF 286
Qy 315 LDREYTLSEKHLKNNMSPCLTLWKAFFNGIOTHSYGVHVFANDITASFTTSKEDMKE 374
Db 287 LVTPHLTVGNNKPKRTEVCVSLVKRVEDVVRDEYAHNFRFTWKTJLSTTIFIS---ETNE 343
Qy 375 FNTTY----HCLNBEIKAEIEKKY-AKVNSTHSGYGLKYFKTDGGLYLWQPLIQNRL 429
XX
```

Db 344 FNLNQIHLSCQKEARAIINRIYTRYNSHVRGTQIOTYLRAGFGVVFQPLLSLSLA 403
 QY 430 -----DAKNKNNETYSRRSQEASTTDDPMWMTGAGGEYSSENSITVAOQVAYDN 484
 Db 404 PLYLQELVRENTNISPQKHPTRNTRSRSSVPELRANT---ITTSVEFAMLQFTYDH 460
 QY 485 LRIRNNILEDLSKAWCREQHRALVWNELSKINPTSVMSIYNRPVSAKRIGDVISVN 544
 Db 461 IQEHVNEMLARISSWCQLQNRERALSGLFPINPSALASTILDORVKARILGDVISVN 520
 QY 545 CIVV-DQTSVLSHLKSLRLLSASDEKCSRPPVPTKFWNDSTIYKGLGVNNEILLTTYL 603
 Db 521 CPELGSDTRIILQNSMR-VSGSTTRCYSRPLISIVSLNGSGTVREGQLGTDNELIMSDLL 579
 QY 604 ETQCENTEYFQAKTDMVIYKNYEHKLTVPPLSSITTLDTFIALNFTLLENVDFKVELYT 663
 Db 580 EPCVANKRYFLFGHHVYEDYRVREIAVHDVGMISTYVDLNLTLKOREFMPLOVYT 639
 QY 664 RDEKRLSNVEDIETMFREYNYAQRVSGLRKDL-LDLSTNRNQFVDAFGSLMDDDLGAVGQ 722
 Db 640 RDELRTDGLLDYSEIQRRNQMHSLRFYDIDKVQYDSGT---AIMQMAQFFQGLGTAGQ 696
 QY 723 TVNNAVSGVATLFSSTIVTGFINFKNPPGGMMLIIVIGULFAIYFLTKTKIYETAPIK 782
 Db 697 AVGHVILGATGALLSTVHGFTTFLSNPFGALAVGLLVLAGLVAFAFFAYRYVYLKLTSPMK 756
 QY 783 MIYPEIDK-LKE-REGKSEIA-----PISE-----EELERIV 812
 Db 757 ALYPLTKGLQKLEGMDPPFAEKNATDTPLEEGDSQNTSPSVNSGDPDPKPREAQEMI 816
 QY 813 LAMHIHQNSHMETKTRDKPKDSILTRAQ-----NNMLKRKSGYSNLK 854
 Db 817 KYMTLVSAABEQESKARKNKTSALLISRLTGLALRNRGYSRV 861

RESULT 34
 AAW34553
 ID AAW34553 standard; Protein; 904 AA.
 XX AC AAW34553;
 AC AAW34553;
 DT 26-FEB-1998 (first entry)
 XX DE
 XX KW Herpes simplex virus type 2 glycoprotein B.
 XX KW Herpes simplex virus; HSV; type I; type II; glycoprotein B; vaccine;
 KW membrane protein; virus-specific glycoprotein;
 XX transmembrane anchor region.
 XX OS Herpes simplex virus type 2.
 XX PN US5648079-A.
 XX PD 15-JUL-1997.
 XX PF 08-DEC-1994; 94US-0351875.
 XX PR 20-OCT-1986; 86US-0921730.
 XX PR 06-APR-1984; 84US-0537784.
 XX PR 20-SEP-1990; 90US-0587179.
 XX PR 21-DEC-1992; 92US-0993415.
 XX PR 18-OCT-1993; 93US-0138717.
 XX PA (CHIR) CHIRON CORP.
 XX PI Burke RL, Pacht C, Valenzuela PDT;
 XX DR WPI; 1997-372022/34.
 XX DR N-PSDB; AAT93651.
 XX PT Vaccines against herpes simplex virus infection - containing
 XX recombinant HSV glycoprotein B

PS Disclosure; Fig 4A-F; 33pp; English.
 XX The present sequence is from the Herpes simplex virus (HSV) type 2
 CC strain 333 and represents a glycoprotein B. HSV is a double stranded
 CC virus packaged within an icosahedral nucleocapsid enveloped within a
 CC membrane. The membrane includes a number of virus-specific
 CC glycoproteins, with glycoprotein B being one of the most abundant.
 CC Glycoprotein B from both HSV type I and type II are cross reactive. New
 CC vaccines comprising recombinantly produced glycosylated glycoprotein B
 CC that has a deletion of at least part of the transmembrane anchor region,
 CC in combination with a carrier and an adjuvant have been produced. The
 CC vaccines are used for immunising humans against HSV (HSV type 1 or
 CC HSV type 2) by vaccination before or after primary infection with HSV.
 XX Sequence 904 AA;
 QY Query Match 18.2%; Score 829.5; DB 18; Length 904;
 Db Best Local Similarity 27.4%; Pred. No. 3e-52;
 QY Matches 243; Conservative 180; Mismatches 363; Indels 101; Gaps 28;
 Db 24 SLSIAETGVTSPTNTATWSTESP--LTGHYGTDDSHGE-RGNENRDSSEQNKNIYGSP 80
 QY 58 ALGAAPTGDPKPKKXKPKNPTPPAPAGDNATVAAGHATLREHLRDIAKENTDANFY--- 114
 Db 81 STEPPYRVCASGVDVRFOTDHCPCDASD-MVHSEGLIILYKONIPFMRVRYKRYKV 139
 QY 115 -----VCPPTGATVVQEPQRCPTREGQNYTEGIAVVFENATPYKFKATMYKDV 168
 Db 140 TTSTV-----YNGIYSDSITNQHTFYKSIPEWE--TEKMDTIYQCFNSLRNTGGNL 189
 QY 169 TVSQVWFGRHYSQFMGIFED-----RAPVPEEVIDKINAKGVCSTAKYVVRNLE 219
 Db 190 LTVVDRDDINMTVFLQPVGV-----TPDVKRYGSPQELYLEPGWFGVSGVRRRTTVN 241
 QY 220 TTAHFRRDDHETDMELKPANAARTSRGWHITDLKYNPVRVEAF-----HRYGTTVN 270
 Db 242 CELMDMFARSNPDPDFVTATGTVEMSPFWS-GEDDH-ENKMKHPWFVSVINNY-KV 298
 QY 271 CIVVEVDARSVYPYDEFVLATGDFVYMSPFYGVREGSHTHTSYAADRFQVDGFFVARDL 330
 Db 299 DYQNRGTVPGLKTRIFLDREBYTSLWEKHLKNMSYCPCLTLWKAFYNGIQTEHSGSVHFA 358
 QY 331 TTARATAP--ATRNLLTTPKFTVANDVWPKRPSVCTMTKWQEVDEMLRAEYGGSFSS 388
 Db 359 NDITASFTT-----SKEDMKEFNTTYHCLNEEIKAEIEKKA-KVNSTHSKYGLK 410
 QY 389 DAISTFTTNLTTEYPLSRVDLGD-----CIGKARDAMDRIFFARRYNATHIIVGQ 442
 Db 411 KTDGGLYLVWQPLIQNRLLDAKNKLNNETYSRRSRQAESTTDDPMWMTGNGAGGEYSSE 470
 QY 443 LANGGLIAYQPLLSNTLAE-----LVVREHLREOSRKPPTPTPPPGASANASVERIK 498
 Db 471 NSITVAQVQVAYDNLRIRINILEDLSKAWCREQHRALVWNELSKINPTSVMSIYNRP 530
 QY 499 SSIEFARLQFTYNIHQHVNDMLGRVAIAWCELQNHLELTWTEARKLNPAINATVGR 558
 Db 531 VSAKRIGDVISVNCIIVDQTSVLSHLKSLRLLSASDEKCSRPPVPTKFWNDSTIYKGL 590
 QY 559 VSARMILGDVNAVSTCVPVAAADNVIVQNSMR-ISSREGACYSRPLVSFRYEDQGPLVEG 617
 Db 591 GVNNEILLTTTLETQCENTEYFQAKTDMVIYKNYEHKLTVPPLSSITTLDTFIALNFTL 650
 QY 618 GENNELRLTRDALEPCTVGHRRYFTFGGGVYVFEESAYSQHSLSRAITTVSTFIDNITM 677
 Db 651 LENVDFKVIELYTRDEKRLSNVFDIETMFREYNYAQRVSGLRKDL-LDLSTNRNQF--VD 708
 QY 678 LEDHEFVPLEVYTRHEIKDSGLLDYTEVQRRNQHLDRFADI-DTVIHADANAAMFAGLG 736
 Db 709 AFGSLMDDDLG-AVGQTVNNAVSGVATLFSSTIVTGFINFKNPPGGMMLIIVIGULFAI 767
 QY 737 AFPEGMDLGRVAVGVGMIGVGVW-----SAVSGVSSFSMSPFPGLAVGLLVLAGL 792
 Db 768 FLTKTKIYETAPIKMIYPEIDK-LK-----EREKSEIAPISE-EELERIVLA 814

XX 19-NOV-1993; 93US-0156866.
 XX PF
 XX 29-OCT-1991; 91US-0788123.
 XX PR
 XX 19-NOV-1993; 93US-0156866.
 XX PR
 XX (UYDE) UNIV DELAWARE.
 XX PA
 XX Keeler CL, Poulsen DJ;
 XX PI
 XX WPI; 1995-302091/39.
 XX DR
 XX N-PSDB; AAQ97351.
 XX XX
 Isolated Infectious Laryngotracheitis Virus gB gene - used to
 PT develop prods. for detection and for vaccine(s), partic. for
 PT immunising fowl.
 XX PT
 XX
 Claim 1; Columns 11-16; lipp; English.
 XX PS
 XX The infectious laryngotracheitis virus (ILV) gB glycoprotein
 CC nucleic acids can be used as probes, for antisense control of gB
 CC gene expression or for production of gB polypeptides. The gB
 CC polypeptides can be used for the production of antibodies (both
 CC used in immunoassays) and in vaccines. Recombinant avipox virus
 CC expressing the ILV gB can be used for immunising fowl, especially
 CC chickens.
 CC (Updated on 25-MAR-2003 to correct PF field.)
 XX CC
 XX Sequence 873 AA;
 SQ
 Query Match 18.1%; Score 821; DB 16; Length 873;
 Best Local Similarity 27.6%; Pred. No. 1.2e-51;
 Matches 220; Conservative 158; Mismatches 310; Indels 108; Gaps 23;
 QY 57 SHGERGNNEARD--SEEQNKIY-----GSPSTFPYR---VCSASGVGDVRFQTDHV 104
 DB 24 SHGIAGIIDPRTASMDVGKISFSEATGSGAPKEPQIRNRFACSSPTGASVARLAQPRH 83
 QY 105 CPDASDMVH-SEGILLIYKQMIIPFMRVRKYRVKVTSTVYNGIYSDSITNQHTFYKSI 163
 DB 84 CHRHADSTNMEGIAVVFKNIAPVFNVTLYKHITVTWALFSPQITNEVTRVPI 143
 QY 164 EPWETEKMDTYTQCFNSLRLNTGNGLITYVDRDINMTVFLOP-----VDGVTPDV 214
 DB 144 DYHEIVRIDRSGECSSKATYHKNFMEFFEAYDNDEREKKLPLVPSLLRSTVSKAPHTTNT 203
 QY 215 KRYGSOPELYLEPCGFWGVSRRRTVNCCLMDMFARSNPDPFVTATGDTVENSFWSG 274
 DB 204 KRHOTL-----GYRTSTSDVCVEYLQARSVPYDYFGMATGDTVEISPFT- 250
 QY 275 EDDHENKMKHPFVSVINNYK---VVDYQNR---GTVPLGKTRIFLDREYVTLSEKH 327
 DB 251 -----KNTGPRRHSVYRDVRFLEIANYQVRDLETQIRPPKKNFLTDEQFTIGHAM 304
 QY 328 LKNMSYCLTLWKAFYNGIQTEHSGSHFVANDITASFTTSKEDMKSFNTT---YHCL- 382
 DB 305 EEKESVCTLSKWIEVPAVRVSYKNSYHFSKDMTMTFFSGKQ---PFNISRHLAECPV 361
 QY 383 ---NEEIKAELEKXKAVNSTHXYGDKYFKTDGGLYLVNVOPLIQNRLLD---AKNKLN 436
 DB 362 TIATEADIGTFARKY---SSTHVRSGDIEYVLGGGFLIAFQKLMGSLAEWYLEEAQRQ 418
 QY 437 NETYSRRSRQAESTTDPMMEMTGNAGGEGYSSENSITVAQVQVAYDNLRIRINILEDL 496
 DB 419 NHLPRGRERRQAAGRRTASLQ--SGPOGDRITHTSSATFAMQLQFAYDKIOAHVNNELGNL 476
 QY 497 SKAWCRQHRALVWNELSKINPTSVMSIYNRPVSASRIGDVISVNSCIVDQTSVSLH 556
 DB 477 LEAWCQLNRQLIIVWHEMKLNPNSLMTSLFGQVPSARLLGDI VAVSKCIEPIENIRMQ 536
 QY 557 KSLRLLSASDEKCFSPRPVTKFM-----NDSTIYKGOLGVNNEILLITTYL 603
 DB 537 QSMR-MFGDPTCYRPLVIRFYRSSSPESQFSANSTENHNDILQLGGEHNEILQGRNLI 595

QY 604 ETCQENTYYFOAKTDMYIYKNYEHKTVPLSSITTLDTIALNFTLLENVDFKVIELYT 663
 DB 596 EPCMINHRRYFLLGENVLLYEDYTFVRQVNASIEBEVSIFINLNATILEDLDFVPVEVYT 655
 QY 664 RDEKRLSNVEDIETMPFREYNYAQR-----VSLRKOLLDLSTNRNQFVDAPGSLMDD 716
 DB 656 REELRDTGLNYDDVVRYQIYNKRFRDIDTVIRGDRDAI-----FRAIADFGNTLGE 710
 QY 717 LG-AGQTVNAVSGVATLFFSSIVTGFINKNPFQGM-----LMIIVVIGVLFAIYFLT 770
 DB 711 VGRALGTVMNTAAAVI-----STVSGIASFLSPFAALGIGIAVVSIIIGLAFKYVMN 766
 QY 771 KTKIYETAPIKMIYP 786
 DB 767 LKSN-----PVQLFP 777
 RESULT 37
 AAP70769
 ID AAP70769 standard; protein; 973 AA.
 XX
 AC AAP70769;
 XX
 DT 25-MAR-2003 (updated)
 DT 11-MAR-1991 (first entry)
 XX
 DE Glycoprotein B of herpes simplex virus type 1.
 XX
 KW Herpes simplex virus; glycoprotein B;
 XX
 OS Herpes simplex virus.
 PH Key Location/Qualifiers
 FT Region 11..26
 FT /label=hydrophobic putative signal
 FT Region 86..88
 FT /label=potential glycosylation site
 FT Region 140..142
 FT /label=potential glycosylation site
 FT Region 429..431
 FT /label=potential glycosylation site
 FT Region 558..560
 FT /label=potential glycosylation site
 FT Region 803..840
 FT /label=anchor region
 FT Region 844..864
 FT /label=anchor region
 FT Region 888..890
 FT /label=potential glycosylation site
 FT Region 957..959
 FT /label=potential glycosylation site
 XX EP236145-A.
 PN
 XX 09-SEP-1987.
 PD
 XX 09-MAR-1987; 87EP-0302001.
 PF
 XX 07-MAR-1986; 86GB-0005646.
 PR 01-SEP-1986; 86GB-0021081.
 PR 16-DEC-1986; 86GB-0029988.
 PR 09-MAR-1987; 87GB-0024676.
 XX
 PA (COGE-) COGENT LTD.
 PA (SMIT/) SMITH G L.
 XX
 PI Smith GL, Barrell BG, Cranage MP;
 XX WPI; 1987-251648/36.
 DR
 XX Prodn. of antigenic, human cytomegalovirus polypeptide(s) - by
 PT expressing recombinant vector, able to induce neutralising antibodies,
 PT

PT useful in vaccines.

PS Disclosure; Fig. 2a-d; 58pp; English.

CC The sequence is that of HSV virus glycoprotein B which shares some structural homology with the human cytomegalovirus gB gene prod. and the herpes simplex virus type 1 glycoprotein.

CC (Updated on 25-MAR-2003 to correct PR field.)

CC (Updated on 25-MAR-2003 to correct PA field.)

XX Sequence 973 AA;

Query Match 18.0%; Score 819.5; DB 8; Length 973;

Best Local Similarity 26.0%; Pred. No. 1.8e-51;

Matches 247; Conservative 181; Mismatches 358; Indels 163; Gaps 29;

QY 28 AETGVTSPPNTATWSTESP--LTGHYGDHDSHGE--RGNENRDSBEQKNYIGSPSTPP 84

DB 61 APTGDTKPKKXKPKKPPPPRAGDNATVAAGHATLRFHLRDIKAENTDANFY----- 113

QY 85 YRVCSAGSGVDVPRFOTDHCVDASD--MVHSEGIILLIYKONIIIPMFVRVRYKRVVTTST 143

DB 114 --VCPPTGATVQVQFQPRCPTRPGQNYTEGIAVFKENIAPYKFKATMYKDVTSQ 171

QY 144 V-----YNGIYSDSIITNOHTFYKSIPEWE--TEKMDTIYQCFSRLNTGMLLTYY 193

DB 172 VMFGHRYSQFMGIFED-----RAPVPFEVIDKINAKGVCSTAKVVRNNLETTAF 222

QY 194 DRDDINMTVFLQPDGV-----TPDVKRYGSOPELYLEPGWFGWYSRRRTTNCCELM 245

DB 223 HRDDHETMELKAPANATRTSRGWHTTDLKYNPSVEAF-----HRYGTTVNCIVE 273

QY 246 DMFARNPPDFRVTATGDTVEMSPFWS--GEDDH--ENKMHKEKPFVSVINNY--KVVDYQN 302

DB 274 EVDARSVPYDEFLVATGDFVMSPFYGYRESHTETSAAADRKFQVDGCFVARDTTKA 333

QY 303 RGTVPGLKTRIFLDREYTLSEKHLKNSYCPLTLWKAFYNGIQTEHSGSYHFVANDIT 362

DB 334 RATAP--TTRNLLTTPEKFTVAMDVWPKRPSVCVTMTKQVEDEMLREYSGSRFSSDAIS 391

QY 363 ASFTT-----SKDMKEFNNTYHCLNEEIKAEIKKYA--KVNSTHSKYGBLKFYKIDG 414

DB 392 TTTFTNLTYPLSRVLDGD-----CIGKDARDMDRIFARRYNATHIKVGPQVYLANG 445

QY 415 GLYLVNQPLIQNLNLL-----DAXKNLNNETYS----- 441

DB 446 GLFIATQPLSNTLAELYVREHLREOSRKPPTPPPGASRDAMDRIFARYNATHIKV 505

QY 442 -----RRSRQAESTTDPMMEMTGNAGGEYS 468

DB 506 GQPVYLANGGLIAQPLSNTLAELYVREHLREOSRKPPTPPPGASANASVERIK 565

QY 469 SENSITVAQVAYDNLRIINILEDKAMCREQRAALVWNLKSNPTSVMSIYN 528

DB 566 TTSISFARLQFTYNNHIOHVNMDLGRVAIACWELQNHETLTWNEARKLNPAIASATVG 625

QY 529 RPSAKRIGDVSVCNIVUDOTSLSLHSLALLSASDEKCFSPRPVTFKFNDSITYKG 588

DB 626 RVSARMUGDMVAVTCVPAADNVIVQNSMR--ISSRPGACYSRPLVSFRYEDQGPLVEG 684

QY 589 QLVGNNELITLTYTEQNTYEFQAKTDMVIYKNEYHLKTVPLSSITLDTFIALNF 648

DB 685 QLVGNNELITRDAIEPCTVGHRRYFTFGGVYVYFEYAYSHQLSRADITTVSTFDLNI 744

QY 649 TLLENVDKVIETYRDEKRLSNVFDIETMFREYNYAQRVSGLRKDLDLSTNRNQF-- 706

DB 745 TMLDEHEFVLEYVYTHEIKDGLDYTEVQRNQLDLRFADI--DTVIHADANAAMEFAG 803

QY 707 VDAFGSLMDDLG--AVGQTVNAVSGVATLFFSSIVTGFINFINKNPGGMLMIIVICVLPFA 765

DB 804 LGAFFEGMDGLGRAVGKVMGIVGGVV-----SAVSGVSSFMGNPFGALVGLVLGLAA 859

QY 766 IYPLTKTKIYETAPIKMIYPEIDK-LK-----EREKSKSIAPISE--EELERIV 812

DB 860 AFFAPRYVMRLQSNPMKALYPLTTTKELKNPTNPASGEGEGGDFDEAKLAAREMIRYM 919

QY 813 LAMHIHQONSHWETKTRKDPKDSILT-PAQNMWML---RKRSGYSNLKNAE 857

DB 920 ALVSAMERTEH---KAKKKGTSAALLSAKVTDVMVRKRRNTNTYTQVPNKD 965

RESULT 38

AAR92746

ID AAR92746 standard; protein; 891 AA.

XX AC AAR92746;

XX DT 11-MAY-1996 (first entry)

XX DE B virus gB glycoprotein.

XX KW Herpes simian monkey B virus gB glycoprotein; UL27; immunoassay;

XX OS Herpes simian monkey B virus.

XX PN US5487969-A.

XX PD 30-JAN-1996.

XX PF 01-APR-1993; 93US-0042747.

XX PR 01-APR-1993; 93US-0042747.

XX PA (SWBI-) SOUTHWEST FOUND BIOMEDICAL RES.

XX PI Black D, Eberle R, Hilliard J, Scinicariello F;

XX DR WPI; 1996-105220/11.

XX PT N-PSDB; AAT16474.

XX PT Detection of herpes B virus by PCR amplification of sample DNA - to

XX PS detect a specific herpes simian monkey B virus DNA segment.

XX SS Claim 1; Column 19-24; 22pp; English.

CC The herpes simian monkey B virus proteins, such as the gB

CC glycoprotein (UL27), have immense potential use in the development

CC of serological immunoassays, which can specifically detect virus

CC antigens and/or antibodies to B virus. One approach is to

CC synthesize peptides which, based on the properties of the predicted

CC protein sequence, are likely to be immunologically active. Such

CC peptides can be used as substrate antigens in immunoassays to detect

CC serum antibodies which recognize this specific peptide sequence.

CC Synthetic peptides may also be used to produce antibodies against

CC specific regions of the gB glycoprotein which are unique to one

CC virus. These can then be used to develop virus-specific

CC immunoassays for differentiation of B virus from other primate

CC alpha-herpes virus and for identification of antibodies directed

CC against B virus in primate serum samples.

XX Sequence 891 AA;

Query Match 18.0%; Score 818.5; DB 17; Length 891;

Best Local Similarity 27.0%; Pred. No. 1.9e-51;

Matches 251; Conservative 169; Mismatches 359; Indels 109; Gaps 26;

QY 2 AGSLKRGSLVIALWYLVQVALYSLS-----IAETGVTSPP-----PNTATWSTESPLTGHY 51

DB 5 AGPLPLPSPLVPLAL---ALLAATRLPLGPAAPVVSPPASPAAPPVPAATPTFPDDND 61

QY 52 GTHUSSGGERGNENRDS-----EEQKNYIGSPSTFPYRVCSAGSGVDVPRFQTDHVC 106

DB 62 GEAGAACGAPGTNASVAGHATLRENLRDIKALDGDATFYVCPPTGATVYVQFQPRCP 121

QY 107 DADSMVHSEGILLIYKQNIIPFMRVRYKRVKVTSTV-----YNGIYSDSIITNOH 157

Qy 204 LQPVdGVTPdVKRYGSQPELYLEPGWFWGSYRRRTTVNCELMDMFARSNPPFFdFVTATG 263

Db 596 RYFLFGGYALFENYFVKWDAADIQIASTFVELNLTLLODREILPLSVVTKRELDRVG 655
Qy 672 VFDIETMREYNYAQRVSLRKDLLDLSTNRNQFVDAFGSLMDDLGAVGQTVVNAVSGV 731
Db 656 VFDYAEVARRNQLHELKFDINK-VIEVDYNY-----AGLOEFGCMSITKKDA---- 702
Qy 732 ATLFSSIVTGFINFIKNPFGG-----MLMIIWIGVLFAIYFLTKTKIYETAPIKM 783
Db 703 -----NPNNGQTOLEAARMELTDLIINAKAMTSLASLODYAKI---EASLSS 744
Qy 784 IYPEIDKIKEREKGEKSEIAPISEEELERIVLAMHIHQONSHMETKTRKDP-----KD 834
Db 745 AYSEAEVTNNLN-----ATLEQLKMAKTNLESAINQANTDKTTDPDNEHPNLVEAYKAKLT 800
Qy 835 SILTRAQNMRLKRS-GYSNLKNAESVEMLN 863
Db 801 TLEQRATNLEGLSTAYNQIRN-NLVDLYN 829

Search completed: October 28, 2003, 15:33:08
Job time : 99 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 28, 2003, 15:31:05 ; Search time 81 Seconds
(without alignments)
1788.314 Million cell updates/sec

Title: US-10-055-364-24

Perfect score: 4547

Sequence: 1 MAGSLKRGSLVLAQWLYQV.....KRGYSNKLKNAESVEMLNTL 865

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 629382 segs, 167460630 residues

Total number of hits satisfying chosen parameters: 629382

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:*

- 1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
- 2: /cgn2_6/ptodata/2/pubpaa/PTCT_NEW_PUB.pep:*
- 3: /cgn2_6/ptodata/2/pubpaa/US05_NEW_PUB.pep:*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep:*
- 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
- 6: /cgn2_6/ptodata/2/pubpaa/PTCTUS_PUBCOMB.pep:*
- 7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
- 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
- 9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep:*
- 10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep:*
- 11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep:*
- 12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep:*
- 13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep:*
- 14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep:*
- 15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep:*
- 16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep:*
- 17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*
- 18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	4547	100.0	865	14	US-10-055-364-24
2	2473	54.4	793	14	US-10-055-364-55
3	2438	53.6	830	14	US-10-055-364-45
4	1940	42.7	824	14	US-10-055-364-43
5	1918	42.2	823	14	US-10-055-364-37
6	1916	42.1	792	14	US-10-055-364-42
7	1902	41.8	808	14	US-10-055-364-38
8	1898.5	41.8	791	14	US-10-055-364-41
9	1871.5	41.2	818	14	US-10-055-364-44
10	1844.5	40.6	829	14	US-10-055-364-46
11	1780	39.1	831	14	US-10-055-364-39
12	1775	39.0	847	14	US-10-055-364-40
13	1128	24.8	900	15	US-10-223-538-2
14	1002	22.0	195	14	US-10-055-364-2
15	844	18.6	904	9	US-09-894-998-18

16	844	18.6	904	12	US-10-200-562-18	Sequence 18, Appl
17	844	18.6	904	12	US-10-237-551-18	Sequence 18, Appl
18	844	18.6	904	15	US-10-121-988-18	Sequence 18, Appl
19	754	16.6	1086	9	US-09-147-052-4	Sequence 4, Appli
20	442	9.7	359	14	US-10-055-364-48	Sequence 48, Appl
21	350	7.7	206	14	US-10-127-733-2	Sequence 2, Appli
22	127	2.8	526	9	US-09-815-242-13696	Sequence 13696, A
23	126.5	2.8	3899	15	US-10-171-311-4	Sequence 4, Appli
24	126.5	2.8	3907	15	US-10-171-311-2	Sequence 8, Appli
25	126.5	2.8	3917	15	US-10-171-311-8	Sequence 2, Appli
26	126.5	2.8	3925	15	US-10-171-311-6	Sequence 6, Appli
27	125	2.7	1009	9	US-09-815-242-12141	Sequence 12141, A
28	124	2.7	996	9	US-09-815-242-5251	Sequence 5251, Ap
29	124	2.7	1565	15	US-10-242-056-59	Sequence 59, Appl
30	123	2.7	792	9	US-09-815-242-12327	Sequence 12327, A
31	120	2.6	748	9	US-09-815-242-12792	Sequence 12792, A
32	119	2.6	2175	11	US-09-978-244A-4	Sequence 4, Appli
33	117.5	2.6	602	12	US-10-032-585-7379	Sequence 7379, Ap
34	117.5	2.6	1025	10	US-09-486-734A-4	Sequence 4, Appli
35	117	2.6	1639	15	US-10-087-464-10	Sequence 10, Appl
36	116.5	2.6	1169	9	US-09-815-242-13448	Sequence 13448, A
37	115.5	2.5	1167	9	US-09-815-242-11522	Sequence 11522, A
38	115	2.5	813	10	US-09-732-180-7	Sequence 7, Appli
39	115	2.5	813	12	US-10-054-399A-29	Sequence 29, Appl
40	115	2.5	844	10	US-09-732-180-2	Sequence 2, Appli
41	115	2.5	844	12	US-10-054-399A-24	Sequence 24, Appl
42	114.5	2.5	1169	9	US-09-815-242-13693	Sequence 13693, A
43	114.5	2.5	2297	15	US-10-245-802-20	Sequence 20, Appl
44	114	2.5	2184	12	US-10-304-095-6	Sequence 6, Appli
45	113.5	2.5	1844	15	US-10-242-056-53	Sequence 53, Appl

ALIGNMENTS

RESULT 1
US-10-055-364-24
; Sequence 24, Application US/10055364
; Publication No. US20020155433A1
; GENERAL INFORMATION:
; APPLICANT: Patience, Clive
; TITLE OF INVENTION: Gamma Herpes Virus DNA and Methods of Use
; FILE REFERENCE: 61750-379
; CURRENT APPLICATION NUMBER: US/10/055.364
; CURRENT FILING DATE: 2002-01-23
; PRIOR APPLICATION NUMBER: US/09/612,204
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: US/60/142,736
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: US/60/168,532
; PRIOR FILING DATE: 1999-12-02
; NUMBER OF SEQ ID NOS: 55
; SOFTWARE: Patent in version 3.0
; SEQ ID NO 24
; LENGTH: 865
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Deduced amino acid sequence of porcine gamma herpesvirus gpB gene
US-10-055-364-24

Query Match	100.0%;	Score 4547;	DB 14;	Length 865;
Best Local Similarity	100.0%;	Pred. NO. 0;		
Matches 865;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	MAGSLKRGSLVLAQWLYQVLSIAETGVTSPPTATWSTESPLTGHYTHDSHGE	60	
Db	1	MAGSLKRGSLVLAQWLYQVLSIAETGVTSPPTATWSTESPLTGHYTHDSHGE	60	
Qy	61	RGNNEENRDESEQNKNIVGSPSTFPYRVCSASGVGDVFRFQTDHVCPCDASDMVHSEGI	120	
Db	61	RGNNEENRDESEQNKNIVGSPSTFPYRVCSASGVGDVFRFQTDHVCPCDASDMVHSEGI	120	

```
Qy 121 YKQNIIPMPRRVRYKRVVTTSTVYNGIYSDSIITNOHTFYKSIIEPWETEKMDTIYQCNS 180
Db 121 YKQNIIPMPRRVRYKRVVTTSTVYNGIYSDSIITNOHTFYKSIIEPWETEKMDTIYQCNS 180
Qy 181 LRLTGGNLLTYVDRDINMTVFLQPDVGVTDPVKRYGSOPELYLEPGWFGSYRRRTTV 240
Db 181 LRLTGGNLLTYVDRDINMTVFLQPDVGVTDPVKRYGSOPELYLEPGWFGSYRRRTTV 240
Qy 241 NCELMDFARSNPPDFPVATGDTVEMSPFWSGEDDHENKMKHFKPVSVINNYKVVDY 300
Db 241 NCELMDFARSNPPDFPVATGDTVEMSPFWSGEDDHENKMKHFKPVSVINNYKVVDY 300
Qy 301 QNRGTVPGLKTRIFLDREEVYLSWEKHLKMSYCPLTLWKAFYNGIOTESHSGSVHFVAND 360
Db 301 QNRGTVPGLKTRIFLDREEVYLSWEKHLKMSYCPLTLWKAFYNGIOTESHSGSVHFVAND 360
Qy 361 ITASFTTSKEDMKEFNTTYHCLNEEIKAEIEKKYAKVNSTHSHKYGDLKYFKTDGGLYLW 420
Db 361 ITASFTTSKEDMKEFNTTYHCLNEEIKAEIEKKYAKVNSTHSHKYGDLKYFKTDGGLYLW 420
Qy 421 QPLIQNRLDANKLNNETYSRRRROAESTTDPMMEMTNGAGGEYSSENSITVAQVY 480
Db 421 QPLIQNRLDANKLNNETYSRRRROAESTTDPMMEMTNGAGGEYSSENSITVAQVY 480
Qy 481 AYDNLRIKNNILDLKAWCREOHRALVWNLKINPTSVMSMIYNNRPVSARKIGDVI 540
Db 481 AYDNLRIKNNILDLKAWCREOHRALVWNLKINPTSVMSMIYNNRPVSARKIGDVI 540
Qy 541 SVSNCIYVDOTSVSLHSLRLLSASDEKCFSRPVPVTFKFWNDSTIYKGQGVNNEILLTT 600
Db 541 SVSNCIYVDOTSVSLHSLRLLSASDEKCFSRPVPVTFKFWNDSTIYKGQGVNNEILLTT 600
Qy 601 TYLETCOENTEYFQAKTDMYIYKNYHLKTVPLSSITTLDTFIALNFTLLENVDFKVE 660
Db 601 TYLETCOENTEYFQAKTDMYIYKNYHLKTVPLSSITTLDTFIALNFTLLENVDFKVE 660
Qy 661 LYTRDEKRLSNVDFIETMFREYNYAQRVSGRLKDLDLSTNRNQFVDAFGSLMDDLGA 720
Db 661 LYTRDEKRLSNVDFIETMFREYNYAQRVSGRLKDLDLSTNRNQFVDAFGSLMDDLGA 720
Qy 721 GQTVNNAVSGVATLFSSIVTGFINFKNPGFGLMIIVIGVLFAIYFLTKTKIYETAP 780
Db 721 GQTVNNAVSGVATLFSSIVTGFINFKNPGFGLMIIVIGVLFAIYFLTKTKIYETAP 780
Qy 781 IKMIYPIBKLEREGKSEIAPISEBELERIVLAMHIHQONSHMETKTRKDPKDSILTRA 840
Db 781 IKMIYPIBKLEREGKSEIAPISEBELERIVLAMHIHQONSHMETKTRKDPKDSILTRA 840
Qy 841 QNMLRKESGYSNLKNAESVEMLNTL 865
Db 841 QNMLRKESGYSNLKNAESVEMLNTL 865

RESULT 2
US-10-055-364-55
; Sequence 55, Application US/10055364
; Publication No. US20020155433A1
; GENERAL INFORMATION:
; APPLICANT: Patience, Clive
; TITLE OF INVENTION: Gamma Herpes Virus DNA and Methods of Use
; FILE REFERENCE: 61750-379
; CURRENT APPLICATION NUMBER: US/10/055,364
; PRIOR FILING DATE: 2002-01-23
; PRIOR APPLICATION NUMBER: US/09/612,204
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: US/60/142,736
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: US/60/168,532
; PRIOR FILING DATE: 1999-12-02
; NUMBER OF SEQ ID NOS: 55
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 55
; LENGTH: 793
```

```
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Portion of Acelaphine herpesvirus.
US-10-055-364-55

Query Match 54.4%; Score 2473; DB 14; Length 793;
Best Local Similarity 57.8%; Pred. No. 9.6e-203;
Matches 468; Conservative 130; Mismatches 169; Indels 46; Gaps 8;

Qy 74 KNIYSPSPFPYRVCASGVDVFRQTQDHCVPDASDMVHSEGLLIYKONIIPFPRVR 133
Db 1 KGHSDPSAFPFRVCASNIGDIFRQTSHSCNTKDKHNEGILLIIPKENVIPYFVKR 60
Qy 134 KYRKVVTSTVYNGIYSDSIITNOHTFYKSIIEPWETEKMDTIYOCFNSLRNTGNLLTYV 193
Db 61 KYRKVVTSTVYNGIYADAVTNOHVFSKSPVYETRRMDTIYQCYNSLDVTGNNLLVYT 120
Qy 194 DRDDINMTVLOPDVGVTDPVKRYGSOPELYLEPGWFGSYRRRTTVNCELMDFARSNP 253
Db 121 DNDGSNMTVDLQPDVGLSNSVRRYHSQPELHAEFGMLLGGYRRRTTVNCEVETDARAVP 180
Qy 254 PFDFVTATGDTVEMSPFW-----SGEDDHENKMKHFKPVSVINNYKVVDYQNRG 304
Db 181 PFYFITNIGDTIEMSPFWKAWNETEFGSEPDR-----LTVAKDYRVVDYKFRG 231
Qy 305 TVPLGKTRIFLDREEVYLSWEKHLKMSYCPLTLWKAFYNGIOTESHSGSVHFVANDITAS 364
Db 232 TPOGHTRIFVDKEEYTLSSWAQOFRNISYCRWAHWSFONAIKTEHGKSLHFVANDITAS 291
Qy 365 FITSKEDMKEFNTTYHCLNEEIKAEIEKKYAKVNSTHSHKYGDLKYFKTDGGLYLWQPLI 424
Db 292 FYTPNTQTRVGLGKHVCLANTIESELKSLAKVNDTHSPNGTAQYLYLTNGGLLLWQPLV 351
Qy 425 QNRLDANKLN-----NETYSRRRROAESTTDPMMEMTNGAGGEYSSENSITV 475
Db 352 QOKLLDAKGLLDVAKKQNTTTTTRRRRQRRSVS-----SGIDDVYTAESTILL 403
Qy 476 AQVQAYDNLRIKNNILDLKAWCREOHRALVWNLKINPTSVMSMIYNNRPVSARK 535
Db 404 TQIQFAYDTLRAQINNVLBELSRACREOHRASLMMNELSKINPTSVMSIYGRVPSAKR 463
Qy 536 IGDVISVNCIYVDOTSVSLHSLRLLSA-SDEKCFSRPVPVTFKFWNDSTIYKGQGVNN 594
Db 464 IGDVISVCHVVDQDSVSLHRSRMRVPGDKTHECYSRPVPVTFKFINDSHLYKGQGVNN 523
Qy 595 EILLTTTLETQOENTEYFQAKTDMYIYKNYHLKTVPLSSITTLDTFIALNFTLLENV 654
Db 524 EILLTTTAVEICHENTEYFQGNMNYFYKNYRHVKTMPVGDAVTLDTFVNLNLIVENI 593
Qy 655 DFKVIELYTRDEKRLSNVDFIETMFREYNYAQRVSGRLKDLDLSTNRNQFVDAFGSLM 714
Db 584 DFOVIELYSREBKRMSTAFDIETMFREYNYTQRTVGLRRDLTDLATNRNQFVDAFGSLM 643
Qy 715 DDLGAVGQTVNNAVSGVATLFSSIVTGFINFKNPGFGLMIIVIGVLFAIYFLTKTK 774
Db 644 DDLGVVGTVLNAVSSVATLFSSIVSGIINFKNPGFGLMLFGLIAAIVTITLLNRKAK 703
Qy 775 IYETAPIKMIYPIBKLEREGKSEIAPISEBELERIVLAMHIHQONSHMETK----- 827
Db 704 RFAQNPVQMI-PDKITITSQREELQVDPISKHELDRLMLAMHDYHASKQPEKQDEEQGS 762
Qy 828 TRKDPKDSILTRAQNMLRKESGYSNLKNAESVE 860
Db 763 TTSGPAD--LNKAKNVLRRAAGYKPLKRTDSFE 793

RESULT 3
US-10-055-364-45
; Sequence 45, Application US/10055364
; Publication No. US20020155433A1
; GENERAL INFORMATION:
; APPLICANT: Patience, Clive
```

; TITLE OF INVENTION: Gamma Herpes Virus DNA and Methods of Use
; FILE REFERENCE: 61750-379
; CURRENT APPLICATION NUMBER: US/10/055,364
; CURRENT FILING DATE: 2002-01-23
; PRIOR APPLICATION NUMBER: US/09/612,204
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: US/60/142,736
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: US/60/168,532
; PRIOR FILING DATE: 1999-12-02
; NUMBER OF SEQ ID NOS: 55
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 45
; LENGTH: 830
; TYPE: PRF
; ORGANISM: Alcelaphine herpesvirus
US-10-055-364-45

Query Match 53.6%; Score 2438; DB 14; Length 830;
Best Local Similarity 57.4%; Pred. No. 1e-199;
Matches 459; Conservative 129; Mismatches 166; Indels 46; Gaps 7;
Qy 64 NENRDESEQN---KNYIGSPSTPPYRVCSASGVDFRQTDHVCPCDASDMVHSEGIILLI 120
Db SNIIRQNNNTAKGIHSDPSAFPFRVCSASNIGDIFRQTSHPSCNPKDKHENEGILLI 101
Qy 121 YKONIIPWFRVRKRVKVVTTSTVNGIYSDSITNOHTFYKSIPEWETEKMDTIYQCENS 180
Db FKENIVPVFVKRVKRVKVVTTSTVNGIYADAVTNOHVFVSKVPIYETRRMDTIYQCENS 161
Qy 181 LRLNTGGNLLTVYDRDDINMTVFLQVPDGVTPDKRYGSOPELYLEPGWFGSVRRRTTV 240
Db LDVTVGNNLLVYTDNGSNMTVDLQPDGLSNVRYHSQPEIHAEPGLGYYRRRTTV 221
Qy 241 NCELMDMFARSNPPDFFWTATGDTVEMSPFW-----SGDDHKNKHEKPFVSV 291
Db NCEVTEDARAVPPRYFITNIGDTIEMSPWKAWNETEFSGEPDRT-----LTV 272
Qy 292 INNYKVVDYQNRGTVPGLKTRIFLDREEVTLSEKHLKNMSYCPDLTKAFVNGIOTBSH 351
Db AKDYRVVDYKFRGTQOQGHTRIFVDKEEYTLSSAQFRNISYCRWAHWKSDNALKTBHG 332
Qy 352 GSYHEVANDITASFTTSKEDKEFNNTYHCLNEEIKABIEKKYAKVNSTHSGYGLKYFK 411
Db KSLHEVANDITASFTYPTQTRVLGKHVCLNNTIESELKSLAKVNDTHSPNGTAQYVL 392
Qy 412 TDGGLYLVWQPLIQNRLLDAKNLN-----NETYSRRSRQAESTTDPMMEMTONG 462
Db TNGGLLVWQPLVQOQKLLDAKGLLDVAKKQNTTTTTTTRSRQRORSVS-----SG 444
Qy 463 AGGEYSSENSITVAQVAYDNLRIIRINNILEDLSKAWCRQHRALVNNELSKINPTSV 522
Db IDDVTAESTILLTQIFAYDTLRAQINNVLSELSRAWCRQHRALVNNELSKINPTSV 504
Qy 523 MSMIYNRPVSAKRIGDIVSVSNCCIYVDQTSVSLHKSRLLSA-SDEKCFSRPPVTFKPMN 581
Db MSSYGRPVSAKRIGDIVSVSHCVVVDQDSVSLHSMRVPGDKTHECYSRPPVTFKFIN 564
Qy 582 DSTIYKQGLGVNNEILLTYYLETQCENTEYFQAQKTONYIKNVEHUKTVPLSSITLTD 641
Db DSHLYKQGLGVNNEILLTAVAEICHEVENTEYFQGGNNNNYFYKNRYRHKTPVPGDVATLD 624
Qy 642 TPIALNFTLENVDFKVELYTRDEKRLSNVFDIETMERENYIAQYSGRKLKOLLDLST 701
Db TPWLNLTLENVDFQVIELYSRESKRMSTAFDIETMERENYIYQYRVTGLRDLTDLAT 684
Qy 702 NRNQVDFAGSLMDLGLAVGQTVNNAVSGVATLFSIVTGFNFTKPNFPGGLMILVIG 761
Db NRNQVDFAGSLMDLGLVGVKTVLNAVSVATLFSIVSGIINFKNFPGGLMFLGLIAA 744
Qy 762 VLFAYFTTKTKIYETAPIKMIYPEIDKLKEREKSEIAPISEBELERIVLAMHIHOON 821
Db VVITVILLNRKAKRAQNPVQMIYDPIKTIITSQREELQVDPISKHELDRIIMLAHMDYHAS 804

Qy 822 SHMETK-----TRKDPKD 834
Db 805 KPESKQDEEQSGTTSQPAD 824
RESULT 4
US-10-055-364-43
; Sequence 43, Application US/10055364
; Publication No. US2002015543A1
; GENERAL INFORMATION:
; APPLICANT: Patience, Clive
; TITLE OF INVENTION: Gamma Herpes Virus DNA and Methods of Use
; FILE REFERENCE: 61750-379
; CURRENT APPLICATION NUMBER: US/10/055,364
; CURRENT FILING DATE: 2002-01-23
; PRIOR APPLICATION NUMBER: US/09/612,204
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: US/60/142,736
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: US/60/168,532
; PRIOR FILING DATE: 1999-12-02
; NUMBER OF SEQ ID NOS: 55
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 43
; LENGTH: 824
; TYPE: PRF
; ORGANISM: Eguine herpesvirus 2
US-10-055-364-43

Query Match 42.7%; Score 1940; DB 14; Length 824;
Best Local Similarity 47.0%; Pred. No. 4.9e-157;
Matches 395; Conservative 141; Mismatches 241; Indels 64; Gaps 16;
Qy 11 VLALWLYQVALYSLSIAETGVTSPNPTATWSTESPTGHYTHDSHGERGNENRDS 70
Db VLCLWCV--AALLCQGAQEVVA-----ETTP-----FATH-----RPEVVAE 46
Qy 71 EQNKNIYGSSTPPYRVCSASGV-GDYFRQTDHVCPCDASDMVHSEGIILLIYKONITPFM 129
Db ENPANPF-----LDFRVCGASPTGGEIIFRPLEESCENPTEDKDHIEGIALIYKNIIVPV 101
Qy 130 FRVKYRKVVTTSTVYNGIYSDSITNOHTFYKSIPEWETEKMDTIYOCFNSLRNTGNL 189
Db FNVKRYKIMTSTIYKGSSEDALTNOHTRSYAVLYEVQVMDHYQCFSAVQVNEGHV 161
Qy 190 LTVYDRDDINMTVFLQVPDGVTPDKRYGSOPELYLEPGWFGSVRRRTTVNCELMDMFA 249
Db NTVYDRDGMNETAFLKPADGLTSSITRYQSOPEVYATPRNLLWSYTRTTTTCNCEVTMSA 221
Qy 250 RSNPPDPFVATGDTVEMSPFWSGEDDHNKHEKPFVSVINNYKVVDYQNRGTVPGL 309
Db RSMKPFPEFFVTSVGDITIEMSPFLKENGTEPEKILKRPHSIOQLKNYAVTKY---GVGLG 277
Qy 310 K---TRIFLDREBYTLSEKHLKNMSYCPDLTKAFYNGIOTBSHSGSYHFPVANDITAS 365
Db QADNATFAFALFDYGLSLSWKATTENSSYCDLLWKGSNAIQTQHSNLSLFIANDITAS 337
Qy 366 TTSKEDKEFNNTYHCLNEEIKABIEKKYAKVNSTHSGYGLKYFKTDGGLYLVWQPLIQ 425
Db STPLEEENFNETFKCIWNNTQBEIOKKLKEVEKTHRPNGTAKVYKTTGNLIYVWQPLIQ 397
Qy 426 NRLLDAKNLN-----ETYSRRSRQAESTTDPMMEMTONGAGGEYS--ENS 472
Db IDLLDTHAKLYNLNTNATASPTSTPTTSPRRRRDTS-----VSGGNGNNGDNSTKEES 450
Qy 473 ITVAQVQYAYDNLRIIRINNILEDLSKAWCRQHRALVNNELSKINPTSVMSMYNRPVS 532
Db VASQVQFAYDNLKRSINRVGELSRAWCRQHRALVNNELSKINPTSVMSAIYGRPV 510
Qy 533 AKRIGDIVSNICIVVDQTSVSLHKSRLLSASDEKCFSRPPVTFKPMNSTIYKQGLV 592
Db AKLIGDVVSDCISVDQKSVFVHKNMK-VPGKEDLCYTRPVVGFKFIINGSELPAQOLGP 569


```

QY 309 GKTRIFLDREBYTLSEWHLKNNMSYCPLTLWKAIFYNGIQTEHSGSYHFVANDITASFTTS 368
Db 266 HKKRIFAHTLDYSVSWEAVNKSASVCSVMFWKSFQRAIQTEHDLTYHFHIANEITAGFSTV 325
QY 369 KEDMKEFNTTYHCLNEEIKAEIEKKYAKVNSTHSHYKGYDKPKYKPTDGGGLVLMQPIQLNRL 428
Db 326 KEPLANFTSDYNLTHINTTLEDKIARVNNHTPNGTAEYQTEGGMLVLMQPIALIEL 385
QY 429 LQAKNKLNE-----TYSRERSQCAESTTDPMMEMTNGAGGEYSSENSITVAQVOYA 481
Db 386 EAMLEAATTPVTPSAPTSRSSKRAIRSIRDV-----SAG-----SENNVFLSQIOYA 434
QY 482 YONLRIRINNILEDISKAWCRQHRRAALVWNLSEKINPTSMVSMIYNRPVSAKRIGDVIS 541
Db 435 YDKLRQSINNVLLEALITWCREQVRQTMVWYIEAKINPTSVMTAIYKRPVSRKALGDVIS 494
QY 542 VENCIVVDOTSVLHKSLLLSASDEKCFSPRPVPTFKFMDNSTIYKQGLGVNNEILLTTT 601
Db 495 VTECINVDOSSVSIHKSLLK--TENNDICYSRPPVTFKFNVSQLEKQGLGARNELLSES 552
QY 602 YLETCOENTYFFOAKTDMYIYKNYEHKLTVPSSITLTDTFIALNFTLLENVDFKVEL 661
Db 553 LVENCHONAEFTTAKNETHYKFNVHVETPLVNNISLTDTFLALNFTFIENIDFAVEL 612
QY 662 YTRDEKRLSNVFDIETMFREYNYAQVSGRLKDLDS--TNRNQFVDAFGSLMDDLGA 719
Db 613 YSSGERKANVFDLEMTFREYNYAQVSGRLKDL--FDNSORNNRDRRIQDFSEILLADLGS 671
QY 720 VQGTVMNAVGVATUFGSSIVTGFINFIRKPNFGGLMIIIVIGVLFALVPLTKTKIYETA 779
Db 672 IGVIVNVASGAFSLFGGIVTGILNFIKPNPLGCMFTLLIGAVIILVILLVRRTNNMQA 731
QY 780 PIKMIYPEIDKLEREGKSEIAPISEEELERIVLAWHITHQONSHMETKTRKDPKOSILTR 839
Db 732 PIRMIYDVEK-----SKSTVTMPBPETIKOILGHNHMQOEAYKKKEQRAARPSIFRQ 786
QY 840 A 840
Db 787 A 787

```

RESULT 7

```

US-10-055-364-38
; Sequence 38, Application US/10055364
; Publication No. US20020155433A1
; GENERAL INFORMATION:

```

```

; APPLICANT: Patience, Clive
; TITLE OF INVENTION: Gamma Herpes Virus DNA and Methods of Use

```

```

; FILE REFERENCE: 61750-379
; CURRENT APPLICATION NUMBER: US/10/055,364

```

```

; CURRENT FILING DATE: 2002-01-23

```

```

; PRIOR APPLICATION NUMBER: US/09/612,204

```

```

; PRIOR FILING DATE: 2000-07-07

```

```

; PRIOR APPLICATION NUMBER: US/60/142,736

```

```

; PRIOR FILING DATE: 1999-07-08

```

```

; PRIOR APPLICATION NUMBER: US/60/168,532

```

```

; PRIOR FILING DATE: 1999-12-02

```

```

; NUMBER OF SEQ ID NOS: 55

```

```

; SOFTWARE: Patent in version 3.0

```

```

; SEQ ID NO 38

```

```

; LENGTH: 808

```

```

; TYPE: PRT

```

```

; ORGANISM: Rhesus monkey rhadinovirus

```

```

US-10-055-364-38

```

```

Query Match 41.8%; Score 1902; DB 14; Length 808;
Best Local Similarity 44.2%; Pred. No. 8 5e-154;
Matches 384; Conservative 142; Mismatches 240; Indels 102; Gaps 14;

```

```

QY 11 VIALMWLYQVALYSIAETGVTSPNTATWSTESPLTGHYTHDSHGERGNENRDS 70
Db 10 LLRAWV--IIAIGTAVGE-NVTTPKGATTAKTP----- 42

```

```

QY 71 BONKNIYGPST-----FPYRVCSAGVGDFRFRQTDHVCDDASDMVHSEGIL 118
Db 43 -----GPSTPTPPENPPRAEAFKFRVCSASATGELFRNLEKTCPTGTEKTHQEGIL 94
QY 119 LIYKONIIPFMRVRYKRVKVVTTSTVYNGIYVSDITNQHTFYKSIETMETSKMDTIYQCF 178
Db 95 MVFKKNIIVPHIFKVRRYKRVKVVTTSTVYNGIYVSDITNQHTFYKSIETMETSKMDTIYQCF 154
QY 179 NSLRINTGNNLLTYVDRDDINMKTVPLOVDGVTDPVKRYGSOPELYLEPGHFWGSRRT 238
Db 155 SSMRVNNGIYVNTYDRDFTNQTVLQPEVGLTNIORFSPQVLYTYTTPGPPGLIYRVT 214
QY 239 TVNCELMDMFARSNPPDFDFVTATGDTVEMSPFWSGE-----DHEHKNMHEKMFVSVI 292
Db 215 TVNCEIVDMIARSAPYSYFVTALGDIVEVSPFCHNDSTCSVAEKTENGLCAR-----VL 269
QY 293 NNYKVVDYQNRGTVPVLPKTRIFLDRBEYTLSEWHLKNNMSYCPLTLWKAIFYNGIQTEHSG 352
Db 270 TNYTMVDPATR--APTETRVFADSGEYTVSWKAEDPKSAVCALTLWKTFFPRAIQTTHEA 327
QY 353 SYHFVANDITASFTTSKEDMKFNTTYHCLNEEIKAEIEKKYAKVNSTHSHYKGYDKPKYKT 412
Db 328 SYHFVANDVTATFTSPUSEVANFTGTYSCLEDEVIQKTLNDTIKKLSODTHVTNGSAQYYKT 387
QY 413 DGGLYLVNQPLIQNRLLDKAKNNETYS-----RRSRQAESTTDPMMEMTGN 461
Db 388 EGGFLNQPLTLPSLDVDEMGNGTTPAPPATTTSTVSRVRSVNTNEQATD----- 439
QY 462 GAGGEYSSENSITVAQVQYAYDNIRININLEDLKAWCRQHRRAALVWNLSEKINPTS 521
Db 440 -----NLAAPQLQFAYDKLRASINKVLEELSRACREQVRDTYMYELSKINPTS 489
QY 522 VMSMIYNRPVSAKIGDVISVNCIVVDOTSVLHKSLLLSASDEKCFSPRPVPTFKPMN 581
Db 490 VMTAIYGRPVSAKIFVGDIAISVTDCVADQASVSIHKSLL--TSTPGMCYSRPPVTFRFLN 547
QY 582 DSTYKQGLGVNNEILLTYYTETCOENTYFFOAKTDMYIYKNYEHKLTVPSSITLTD 641
Db 548 STTLFKQGLGPRNEIILTDNQVEACKETCEHYFTASNVTYKYDYVFKKINTSEISLTG 607
QY 642 TFIALNFTLLENVDFKVIETYTRDEKRLS--NVFDIETMFREYNYAQVSGRLKDL---L 697
Db 608 TFIALNLSFIENIDFRVIELYSRAEKLSGSVFDIETMFREYNYAQVSGRLKDL---L 667
QY 698 DLSTNRNQFVDAFGSLMDDLGAQGTVMNAVGVATUFGSSIVTGFINFIRKPNFGGLMII 757
Db 668 DL--NRDRLARDLSEIVADLGDVGRTVMNAVSSVITLFGSIVSGFINFIRKPNFGGLMII 725
QY 758 VVIGVLFALVPLTKTKIYETAPIKMIYPEIDKLEREGKSEIAPISEEELERIVLAWHI 817
Db 726 VIVAVLVFALNRRTNAIAQAPIRMIVPDIKMQPSGGK-----VDOEQIKNILAGMHQ 780
QY 818 HQONSHMETKTRKDPKOSILTRAQNMRL 845
Db 781 LQEBERRRLDEQORSAPSLPRASDGLK 808

```

RESULT 8

```

US-10-055-364-41

```

```

; Sequence 41, Application US/10055364

```

```

; Publication No. US20020155433A1

```

```

; GENERAL INFORMATION:

```

```

; APPLICANT: Patience, Clive

```

```

; TITLE OF INVENTION: Gamma Herpes Virus DNA and Methods of Use

```

```

; FILE REFERENCE: 61750-379

```

```

; CURRENT APPLICATION NUMBER: US/10/055,364

```

```

; CURRENT FILING DATE: 2002-01-23

```

```

; PRIOR APPLICATION NUMBER: US/09/612,204

```

```

; PRIOR FILING DATE: 2000-07-07

```

```

; PRIOR APPLICATION NUMBER: US/60/142,736

```

```

; PRIOR FILING DATE: 1999-07-08

```

```

; PRIOR APPLICATION NUMBER: US/60/168,532

```

```
; PRIOR FILING DATE: 1999-12-02
; NUMBER OF SEQ ID NOS: 55
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 41
; LENGTH: 791
; TYPE: PRT
; ORGANISM: Ateline herpesvirus 3
US-10-055-364-41

Query Match      41.8%; Score 1898.5; DB 14; Length 791;
Best Local Similarity 46.6%; Pred. No. 1.6e-153;
Matches 367; Conservative 144; Mismatches 235; Indels 41; Gaps 10;

Qy 69 SEQKNYIGSPSPFYRVCASGVDPFRFOTDHCPCDASDMVHSEGLLIYKONIIPF 128
Db 26 SDENKTPAIEKEFKYRVCASASTGELFRNLDRACPSTEDKVRREGILLVYKKNIVPH 85

Qy 129 MFRYKRVKVVTTSTVNGIYSD--SITNQHFTFKYSIEPMEETERKMDTIYQCFNSRLNTG 186
Db 86 IFKVRVYKKIATSVRIFGWMSREGVAITNKWELGRAVPKYEINLMDKNYQCHNCMQIEVN 145

Qy 187 GNLLTYVDRDDINMTVFPLOPDVGTDPVKRYGSOPELYLEPCWFWGYSYRRRTTVCNCELD 246
Db 146 GLLSYCDRDGNNTKVDLPKPDGDTGATRYVSQPKIFADAGMLWGTGKTRTTVNCNCEIVE 205

Qy 247 MFARSNPPDFPVATGDTVEMSPFWSGEDDHENKHEKMPFVSVIN----NYKVVDYQ 301
Db 206 MFARSADPYTFVVALGDTVEVSFPCDAENSCPNAS-----DVLSSQVDFNHTVVVDYG 258

Qy 302 NRGTVPLGKTRIFLDREBYTLWSKHLKNMSYCPCLTLWKAFYNGIQTEHSGSYHFVANDI 361
Db 259 NRATSQOHGKRIFAHTLDYSVWEAINKTTSVCMSVFWKGFQRAIQTEHSDTYHFIANEI 318

Qy 362 TASTTSKEDMKNENTYHCHLNEBIKAEIEKKYAKVNSTHSHKYGDLYKFTDGLGLYLWQ 421
Db 319 TAGFESTKETLASFSESYSCILMSDINSTLTDKIGRVNNTHPNGTAQYFXTGEGMILWQ 378

Qy 422 PLIQNRLLDANKLNN-----EYSSRSRRAESTTDPMMEMTGCAGGSEYSENSITV 475
Db 379 PLTAIELEAMEIATTVSPPLSTAHLTSTRGKRKDV-----SAG-----SENSVLL 427

Qy 476 AQOYAYDNLRIINNILEDLSKAWCREQHRALVNNELSKINPTSVMMSIYRNPVSAGR 535
Db 428 AQIQYAYDKLQSQINNVLLELAITWCREQVQRTMIWYIEAKINPTSVMTAIYKPVSAKA 487

Qy 536 IGDVISVNCIVDQTSVLSHKSURLLSASDEKCFSPPTVTFKPMNDSTIYKQGLGVNNE 595
Db 488 LGDVISVTECINVDQTSVSIHKSILK--TTNNDVCYSRPPVTFKPVNSQLFKGOLGARNE 545

Qy 596 ILLFTTYLETCEOENTEYVFOAKTDMYLYKNVEHLKTVPLSSITTLDTFIALNFTLLENVD 655
Db 546 ILLSESIVENCHQNAEHFFAKNETHYFNKNVLAHVETLPLTNISTITLFLALNFTFENID 605

Qy 656 FKVIELYTRDEKRLSNVDFDIETMFREYNYYAQRSGLRKDLDS--TNRNQFVDFAGSL 713
Db 606 FKAVELYSSEGERKLANVDFLETMFREYNYYAQRSGLRKD--FDNSQRNNDRRIIQDFSEI 664

Qy 714 MDDLCAVGQTVNAVSGVATLFFSSIVTGFNFINKNPFGGMLMIIVUGLFAIYFLTKYT 773
Db 665 LADLGSGIKGVINVIASSAFSLFGGIVTGILNFIKNPLGGMILTFLLVGAIIILVILVRRT 724

Qy 774 KIVETAPIKMYPIDKLEREGKSEIAPISSEELERIIVLAMIHTHOONSHMETKTRDKP 833
Db 725 NNNWSQAFIRMIYDPDIEK-----SRSSVTPTPEPEVIKQILLGMNMQOEEYKKEBHKASQ 779

Qy 834 DSILTRA 840
Db 780 PSFLKRA 786

RESULT 9
US-10-055-364-44
; Sequence 44, Application US/10055364
```

```
; Publication No. US20020155433A1
; GENERAL INFORMATION:
; APPLICANT: Patience, Clive
; TITLE OF INVENTION: Gamma Herpes Virus DNA and Methods of Use
; FILE REFERENCE: 61750-379
; CURRENT APPLICATION NUMBER: US/10/055,364
; CURRENT FILING DATE: 2002-01-23
; PRIOR APPLICATION NUMBER: US/09/612,204
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: US/60/142,736
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: US/60/168,532
; PRIOR FILING DATE: 1999-12-02
; NUMBER OF SEQ ID NOS: 55
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 44
; LENGTH: 818
; TYPE: PRT
; ORGANISM: Equine herpesvirus 5
US-10-055-364-44

Query Match      41.2%; Score 1871.5; DB 14; Length 818;
Best Local Similarity 46.2%; Pred. No. 3.6e-151;
Matches 368; Conservative 149; Mismatches 223; Indels 57; Gaps 13;

Qy 85 YRYCASGVGDVFRFOTDHCPCDASDMVHSEGLLIYKONIIPFMRVRYKRVKVVTTSTV 144
Db 56 FRVCGASPGEIFRFRPLEENCNPTEDEKHEVGEILLIYKTNIVPIYFNVRKYRKLVTSTTI 115

Qy 145 YNGIYSDSITNQHFTFKYSIEPMEETERKMDTIYQCFNSRLNTGMLLTYYVDRDDINMTVFL 204
Db 116 YKWSQDAITNOYTSFSAFPLMEARLVDYNYECYNGIQTEHGLTIYVDRDGYNESVRL 175

Qy 205 QPVDGVTDPVKRYGSOPELYLEPCWFWGYSYRRRTTVCNCELDMFARSNPPDFPVATG 264
Db 176 VPADGLTTSIRRYHSGOPELYVTPRNLMSVYTRTTVNCVCEVIDMTARSHKPPFEYFVATSG 235

Qy 265 TVEMSPFWSGEDDHENKHEKMPFVSVINNYKVVDYQNRGCTVPLGK----TRIFLOREEY 320
Db 236 SIETSPFYT-----NASRRVP--VQVLYNSVTDY-----GVGLSGSENVTFEATLNDP 283

Qy 321 TLSWEKHLKNMSYCPCLTLWKAFYNGIQTEHSGSYHFVANDITASFTTS--KEDMKENFTTY 379
Db 284 SISKKAATENSSYCPVLVWKGPSPAIQTKHEKSYHFIADAVTASFTTPTLTDTSYENTTY 343

Qy 380 HCLNEBIKAEIEKKYAKVNSTHSHKYGDLYKFTDGLGLYLWQPLIQNRLLDANKLNN-- 437
Db 344 QCAWQDIEGIEIQKRPDPVSKTHARNGSVQIYKTSGNLYVWQPLVQLDLAAHAKTINST 403

Qy 438 -----EYSSRSRRAESTTDPMMEMTGCAGGSEYSENSITVAVQVAYD 483
Db 404 DNSTSTPTAPNTTSTSSRRKRDRDTCNTA-----TNSSNNSSMBEELATSSQVQFAYD 457

Qy 484 NLRIRINNILEDLSKAWCREQHRALVNNELSKINPTSVMMSIYRNPVSAGRIGDVISVS 543
Db 458 QLRKSNRVLEQLSRVWCNQYRASLWYELSKINPTSVMSAIYGRPVSAKLGVDDVQVQS 517

Qy 544 NCIVDQTSVLSHKSURLLSASDEKCFSPPTVTFKPMNDSTIYKQGLGVNNEILLTITTYL 603
Db 518 DCITVDQESVFPVHRNLR--VPGSKDLCYTRFPVVGKFEINGSSELPVVGQLGARNEILLSTNLV 576

Qy 604 ETCOENTEYVFOAKTDMYLYKNVEHLKTVPLSSITTLDTFIALNFTLLENVDKVIELYT 663
Db 577 EVCQHSCEHYFOGNNHIYKKNVEYVSTWNLDVDFLHTMTITLNLSLVENVDQVQLQYS 636

Qy 664 RDEKRLSNVDFDIETMFREYNYYAQRSGLRKDLDS--LSTNRNQFVDAFGLSMDDLGAVGQ 722
Db 637 QREKRLSNVDFDIETMFREYNYYAQRSGLRKDLDS--LSTNRNQFVDAFGLSMDDLGAVGQ 696

Qy 723 TVNAVSGVATLFFSSIVTGFNFINKNPFGGMLMIIVUGLFAIY--FLTKTKIYETAPI 781
Db 697 VIVNAVSGVFLFGSIVSGVISFLKNPLGAILTIALIVGGIIVLVILFTRSRVYQ-API 755
```



```
Db 633 GHRRYIFGGVYFEEYAYSHQLSRADVTTVSTFIDLNTLMLEDHEFFVLEVYTRHEIK 692
QY 669 LSNVFDIETMFREYNYAQRVSLGRKDLDDLSTNRNQFVD--AFGLMDDLG-AGQTVV 725
Db 693 DSGLLDYTEVQRNQLHDLRFADI-DTVIRADANAAMFAGLCAFEFGMDLGRVGVVM 751
QY 726 NAVSGVATLFSIVTGFNFIKNPFGGMLIIVIGVLFAIYFLTKTKIYETAPIKMIY 785
Db 752 GVVGGVV---SAVSGVSFMSNPFAGALVGLLVAGLVAFAFPAFYVLQLRNPKALY 807
QY 786 PEIDKLKEREGKSEIAPISSE-----ELERIVLAMHIHQONSHMETKTRKDP 832
Db 808 PLTTKELKTSDPGGVGGEGEAGGDFEAKLAEMIRYMALVSAMERTEHKARKKG 867
QY 833 KDSIL-TRAQNM-LRKR--SGYSNLKNAE 857
Db 868 TSALLSSKVTNMVLRKRNKARYSPLHND 896

RESULT 17
US-10-237-551-18
; Sequence 18, Application US/10237551
; Publication No. US20030165820A1
; GENERAL INFORMATION:
; APPLICANT: Day, Craig H.
; APPLICANT: Hosken, Nancy A.
; APPLICANT: Parsons, Joseph M.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DIAGNOSIS AND
; TITLE OF INVENTION: TREATMENT OF HERPES SIMPLEX VIRUS INFECTION
; FILE REFERENCE: 210121.538C3
; CURRENT APPLICATION NUMBER: US/10/237,551
; CURRENT FILING DATE: 2002-09-06
; NUMBER OF SEQ ID NOS: 254
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 18
; LENGTH: 904
; TYPE: PRT
; ORGANISM: Herpes simplex virus
US-10-237-551-18

Query Match 18.6%; Score 844; DB 12; Length 904;
Best Local Similarity 28.4%; Pred. No. 4.8e-63;
Matches 247; Conservative 164; Mismatches 366; Indels 92; Gaps 27;

QY 33 TSPNTATWSTESPLTGHYTHDSHGERGNENRDSEEQNKNIYGSPTFPYRVCASG 92
Db 76 TTPPD-----ANATVAAGHATLR--AHLREIKVENADAO-----FYVCPPT 115
QY 93 VGDVFRFQTDHVCPSDASD-MVHSEGILLIYKQNIIPFMFRVRKRVKVVTTSTV----- 144
Db 116 GATVVQFEQPRRCPTREPCQNYTEGIAVVFKEKNIAPYKFKATMYKDVTSQVWFGHYS 175
QY 145 -YNGIYSDSTINQHTFYKSIETPWE--TEKMDTIYOCFNSLRNTGNNLLTYVDRDINMT 201
Db 176 QFMGIFED-----RAPVFEVIDKINTKGVCRCSTAKYVRNMMETTAFAHRDDHETD 226
QY 202 VFLQVQDGVTPDKRYGSOPELYLSPGFWGYSRRRTVNCLEMDMFAEARSPPDFVFTA 261
Db 227 MELKPAK-VATRTSGWHTDLDKYNPSVEAPHRVGTTCVNCIIEVDARSVYVDEFVLA 285
QY 262 TGDVTVMSPFWS-GBDDH-ENKMHKPKMFSVINYNY-KVVDYQNRGTVPFGKTRIFLDRE 318
Db 286 TGDVFMSPFYREGSHTSHTSYAADRKFQVDGFGYARDLTITKARATSP--TTRNLLTTP 343
QY 319 EYTLSEKHLKMSYCPILTKAFYNGIQTEHSGSYHYFVANDITASFTT-----SKED 371
Db 344 KFTVAMDVPKRPVACTWTKMOEVDMLRAEYGGGFRFSSDAISTTFTNLTEYSLSRVD 403
QY 372 MKEFNTVHCLNEEBIKAEITEKKYA-KVNSTHSGYKDLXFKYKTDGGLVWQPLONRLLD 430
Db 404 LGD-----CIGRDAREIDRFARKYNATHIKVGPQYIATGCGFLIAYOPLLSNTLAE 457
QY 431 AKNKLNNETYSR-RSRROAESTTDPMMEM-TGNGAGGEYSSENSITVAQVQVAYDNLIR 488
```

```
Db 458 ---LYVREYMQDRKPRNATPAFLREAPASANASVERIKTSSIEFARLQFTYHNQHR 513
QY 489 INNTEIDLSKAWCREQHRALVWNELSKINTPSVMSMIYNRPVSAKRIGDVISVNCIYV 548
Db 514 VNDMLGR:IAVAMCELQNHLETLWNEARKLNPNATASATVGRRVRSARMLGDVNAVSTCPV 573
QY 549 DQTSVLSHKSRLLSASDEKCFSRPPTFKFMDNSTIYKQLGQVNVNEILLTTTYLETQOE 608
Db 574 APDNVIVQNSWR-VSSRPCTCYSRPLVSFRVEDQGPLIEGQCGENNELRLTRDALEPCTV 632
QY 609 NTEYIFQAKTDWYIYKNYEHLKTVPLSSITLDTFIALNFTLLENVDFKVIELYTRDEKR 668
Db 633 GHRRYIFGGVYFEEYAYSHQLSRADVTTVSTFIDLNTLMLEDHEFFVLEVYTRHEIK 692
QY 669 LSNVFDIETMFREYNYAQRVSLGRKDLDDLSTNRNQFVD--AFGLMDDLG-AGQTVV 725
Db 693 DSGLLDYTEVQRNQLHDLRFADI-DTVIRADANAAMFAGLCAFEFGMDLGRVGVVM 751
QY 726 NAVSGVATLFSIVTGFNFIKNPFGGMLIIVIGVLFAIYFLTKTKIYETAPIKMIY 785
Db 752 GVVGGVV---SAVSGVSFMSNPFAGALVGLLVAGLVAFAFPAFYVLQLRNPKALY 807
QY 786 PEIDKLKEREGKSEIAPISSE-----ELERIVLAMHIHQONSHMETKTRKDP 832
Db 808 PLTTKELKTSDPGGVGGEGEAGGDFEAKLAEMIRYMALVSAMERTEHKARKKG 867
QY 833 KDSIL-TRAQNM-LRKR--SGYSNLKNAE 857
Db 868 TSALLSSKVTNMVLRKRNKARYSPLHND 896

RESULT 18
US-10-121-988-18
; Sequence 18, Application US/10121988
; Publication No. US20030068327A1
; GENERAL INFORMATION:
; APPLICANT: Hosken, Nancy Ann
; APPLICANT: McGowan, Patrick
; APPLICANT: Sleath, Paul R.
; APPLICANT: Mossman, Sally P.
; APPLICANT: Evans, Lawrence S.
; APPLICANT: Swanson, Ryan M.
; APPLICANT: McNeill, Patricia D.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DIAGNOSIS AND
; TITLE OF INVENTION: TREATMENT OF HERPES SIMPLEX VIRUS INFECTION
; FILE REFERENCE: 210121.538C1
; CURRENT APPLICATION NUMBER: US/10/121,988
; CURRENT FILING DATE: 2002-04-11
; NUMBER OF SEQ ID NOS: 183
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 18
; LENGTH: 904
; TYPE: PRT
; ORGANISM: Herpes simplex virus
US-10-121-988-18

Query Match 18.6%; Score 844; DB 15; Length 904;
Best Local Similarity 28.4%; Pred. No. 4.8e-63;
Matches 247; Conservative 164; Mismatches 366; Indels 92; Gaps 27;

QY 33 TSPNTATWSTESPLTGHYTHDSHGERGNENRDSEEQNKNIYGSPTFPYRVCASG 92
Db 76 TTPPD-----ANATVAAGHATLR--AHLREIKVENADAO-----FYVCPPT 115
QY 93 VGDVFRFQTDHVCPSDASD-MVHSEGILLIYKQNIIPFMFRVRKRVKVVTTSTV----- 144
Db 116 GATVVQFEQPRRCPTREPCQNYTEGIAVVFKEKNIAPYKFKATMYKDVTSQVWFGHYS 175
QY 145 -YNGIYSDSTINQHTFYKSIETPWE--TEKMDTIYOCFNSLRNTGNNLLTYVDRDINMT 201
Db 176 QFMGIFED-----RAPVFEVIDKINTKGVCRCSTAKYVRNMMETTAFAHRDDHETD 226
```



```

; FILE REFERENCE: 61750-379
; CURRENT APPLICATION NUMBER: US/10/055,364
; CURRENT FILING DATE: 2002-01-23
; PRIOR FILING DATE: 2000-07-07
; PRIOR FILING DATE: 2000-07-07
; PRIOR FILING DATE: 1999-07-08
; PRIOR FILING DATE: 1999-07-08
; PRIOR FILING DATE: 1999-12-02
; NUMBER OF SEQ ID NOS: 55
; SOFTWARE: Patent in version 3.0
; SEQ ID NO 48
; LENGTH: 359
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Suid herpesvirus 1 - bases 491-850
US-10-055-364-48

Query Match
Best Local Similarity 9.7%; Score 442; DB 14; Length 359;
Matches 108; Conservative 83; Mismatches 152; Indels 24; Gaps 8;

; 441 SRRSRQ--AESTDPMWNTGNGAGGESSNSITVAQVQYAYNLIRINNIILEDLSK 498
; 6 ARRARESPGAGTPEP---PAVNGTGHRLITTSAGFAERLQFTYDHIQAHVNDMLGRTAA 62
; 499 AWCREQRAALVNNELSKINPTSVMSIYNRPVSARKIGDVISVNCNIWDQTSVSLHKS 558
; 63 ANCELQNKORTLWSENRLNPAVATAALQQRVSRMLGDVNAISRCVEV-RGVYVQNS 121
; 559 LRLLSASDEKCFSPRPVTFKFNMDSTIYKQGLGVNNEILLTTLTETCOENTYFFQAKT 618
; 122 MR-VPGERGTCSYRPLVTFEH-NGTVIEQGLDGNELLISRDLEPCGTGNHRRYFKLGS 179
; 619 DMWIKYKNEHLKTVPLSSITLDTFTALNFTLLENVDFKVIYLRDEKRLSNWFDIETM 678
; 180 GYVYVEDYNYVRMVEYPE--TISTRTVLTNLTLEDREFLEPLFVYTRTELADTGLLDYSEI 237
; 679 FREYNYVA-----QRVSLGRKDLDDLSTNRNQFVDAFGLSMDLGGVQTVNNAVSQA 732
; 238 QRRNQHLAKFYDIDRVKVDNVLLRGIANFF-----QGLGDVGAAGVKVVLGAT 289
; 733 TLFSSIVTGFNFINKPFGGMLMIIVVIGVLFAYFLTKTKIYETAPIKMIYPIDKLK 792
; 290 GAVISAVGMVSLSPNFGALGLVLGLVAGLVAFLAYRHISRLRNPMKALYPVTTKL 349
; 793 EREGKSE 799
; 350 KEDGVDE 356

RESULT 21
US-10-127-733-2
; Sequence 2, Application US/10127733
; Publication No. US20020150889A1
; GENERAL INFORMATION:
; APPLICANT: EICKMANN, Markus
; APPLICANT: GICKLHORN, Dorothee
; APPLICANT: RADSAK, Klaus
; APPLICANT: HAUSER, Hans-Peter
; APPLICANT: GIESENDORF, Bernhard
; TITLE OF INVENTION: IMMUNOREACTIVE REGIONS OF GLYCOPROTEIN GP11 OF
; TITLE OF INVENTION: VARICELLA ZOSTER VIRUS (VZV)
; FILE REFERENCE: 058315/0128
; CURRENT APPLICATION NUMBER: US/10/127,733
; CURRENT FILING DATE: 2002-04-23
; PRIOR APPLICATION NUMBER: US/09/219,797
; PRIOR FILING DATE: 1998-12-23
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: Patent in Ver. 2.0
; SEQ ID NO 2
; LENGTH: 206

; FILE REFERENCE: 61750-379
; CURRENT APPLICATION NUMBER: US/10/055,364
; CURRENT FILING DATE: 2002-01-23
; PRIOR FILING DATE: 2000-07-07
; PRIOR FILING DATE: 2000-07-07
; PRIOR FILING DATE: 1999-07-08
; PRIOR FILING DATE: 1999-07-08
; PRIOR FILING DATE: 1999-12-02
; NUMBER OF SEQ ID NOS: 55
; SOFTWARE: Patent in version 3.0
; SEQ ID NO 48
; LENGTH: 359
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Suid herpesvirus 1 - bases 491-850
US-10-055-364-48

Query Match
Best Local Similarity 9.7%; Score 442; DB 14; Length 359;
Matches 108; Conservative 83; Mismatches 152; Indels 24; Gaps 8;

; 441 SRRSRQ--AESTDPMWNTGNGAGGESSNSITVAQVQYAYNLIRINNIILEDLSK 498
; 6 ARRARESPGAGTPEP---PAVNGTGHRLITTSAGFAERLQFTYDHIQAHVNDMLGRTAA 62
; 499 AWCREQRAALVNNELSKINPTSVMSIYNRPVSARKIGDVISVNCNIWDQTSVSLHKS 558
; 63 ANCELQNKORTLWSENRLNPAVATAALQQRVSRMLGDVNAISRCVEV-RGVYVQNS 121
; 559 LRLLSASDEKCFSPRPVTFKFNMDSTIYKQGLGVNNEILLTTLTETCOENTYFFQAKT 618
; 122 MR-VPGERGTCSYRPLVTFEH-NGTVIEQGLDGNELLISRDLEPCGTGNHRRYFKLGS 179
; 619 DMWIKYKNEHLKTVPLSSITLDTFTALNFTLLENVDFKVIYLRDEKRLSNWFDIETM 678
; 180 GYVYVEDYNYVRMVEYPE--TISTRTVLTNLTLEDREFLEPLFVYTRTELADTGLLDYSEI 237
; 679 FREYNYVA-----QRVSLGRKDLDDLSTNRNQFVDAFGLSMDLGGVQTVNNAVSQA 732
; 238 QRRNQHLAKFYDIDRVKVDNVLLRGIANFF-----QGLGDVGAAGVKVVLGAT 289
; 733 TLFSSIVTGFNFINKPFGGMLMIIVVIGVLFAYFLTKTKIYETAPIKMIYPIDKLK 792
; 290 GAVISAVGMVSLSPNFGALGLVLGLVAGLVAFLAYRHISRLRNPMKALYPVTTKL 349
; 793 EREGKSE 799
; 350 KEDGVDE 356

RESULT 21
US-10-127-733-2
; Sequence 2, Application US/10127733
; Publication No. US20020150889A1
; GENERAL INFORMATION:
; APPLICANT: EICKMANN, Markus
; APPLICANT: GICKLHORN, Dorothee
; APPLICANT: RADSAK, Klaus
; APPLICANT: HAUSER, Hans-Peter
; APPLICANT: GIESENDORF, Bernhard
; TITLE OF INVENTION: IMMUNOREACTIVE REGIONS OF GLYCOPROTEIN GP11 OF
; TITLE OF INVENTION: VARICELLA ZOSTER VIRUS (VZV)
; FILE REFERENCE: 058315/0128
; CURRENT APPLICATION NUMBER: US/10/127,733
; CURRENT FILING DATE: 2002-04-23
; PRIOR APPLICATION NUMBER: US/09/219,797
; PRIOR FILING DATE: 1998-12-23
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: Patent in Ver. 2.0
; SEQ ID NO 2
; LENGTH: 206

; FILE REFERENCE: 61750-379
; CURRENT APPLICATION NUMBER: US/10/055,364
; CURRENT FILING DATE: 2002-01-23
; PRIOR FILING DATE: 2000-07-07
; PRIOR FILING DATE: 2000-07-07
; PRIOR FILING DATE: 1999-07-08
; PRIOR FILING DATE: 1999-07-08
; PRIOR FILING DATE: 1999-12-02
; NUMBER OF SEQ ID NOS: 55
; SOFTWARE: Patent in version 3.0
; SEQ ID NO 48
; LENGTH: 359
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Suid herpesvirus 1 - bases 491-850
US-10-055-364-48

Query Match
Best Local Similarity 9.7%; Score 442; DB 14; Length 359;
Matches 108; Conservative 83; Mismatches 152; Indels 24; Gaps 8;

; 441 SRRSRQ--AESTDPMWNTGNGAGGESSNSITVAQVQYAYNLIRINNIILEDLSK 498
; 6 ARRARESPGAGTPEP---PAVNGTGHRLITTSAGFAERLQFTYDHIQAHVNDMLGRTAA 62
; 499 AWCREQRAALVNNELSKINPTSVMSIYNRPVSARKIGDVISVNCNIWDQTSVSLHKS 558
; 63 ANCELQNKORTLWSENRLNPAVATAALQQRVSRMLGDVNAISRCVEV-RGVYVQNS 121
; 559 LRLLSASDEKCFSPRPVTFKFNMDSTIYKQGLGVNNEILLTTLTETCOENTYFFQAKT 618
; 122 MR-VPGERGTCSYRPLVTFEH-NGTVIEQGLDGNELLISRDLEPCGTGNHRRYFKLGS 179
; 619 DMWIKYKNEHLKTVPLSSITLDTFTALNFTLLENVDFKVIYLRDEKRLSNWFDIETM 678
; 180 GYVYVEDYNYVRMVEYPE--TISTRTVLTNLTLEDREFLEPLFVYTRTELADTGLLDYSEI 237
; 679 FREYNYVA-----QRVSLGRKDLDDLSTNRNQFVDAFGLSMDLGGVQTVNNAVSQA 732
; 238 QRRNQHLAKFYDIDRVKVDNVLLRGIANFF-----QGLGDVGAAGVKVVLGAT 289
; 733 TLFSSIVTGFNFINKPFGGMLMIIVVIGVLFAYFLTKTKIYETAPIKMIYPIDKLK 792
; 290 GAVISAVGMVSLSPNFGALGLVLGLVAGLVAFLAYRHISRLRNPMKALYPVTTKL 349
; 793 EREGKSE 799
; 350 KEDGVDE 356

RESULT 21
US-10-127-733-2
; Sequence 2, Application US/10127733
; Publication No. US20020150889A1
; GENERAL INFORMATION:
; APPLICANT: EICKMANN, Markus
; APPLICANT: GICKLHORN, Dorothee
; APPLICANT: RADSAK, Klaus
; APPLICANT: HAUSER, Hans-Peter
; APPLICANT: GIESENDORF, Bernhard
; TITLE OF INVENTION: IMMUNOREACTIVE REGIONS OF GLYCOPROTEIN GP11 OF
; TITLE OF INVENTION: VARICELLA ZOSTER VIRUS (VZV)
; FILE REFERENCE: 058315/0128
; CURRENT APPLICATION NUMBER: US/10/127,733
; CURRENT FILING DATE: 2002-04-23
; PRIOR APPLICATION NUMBER: US/09/219,797
; PRIOR FILING DATE: 1998-12-23
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: Patent in Ver. 2.0
; SEQ ID NO 2
; LENGTH: 206

; FILE REFERENCE: 61750-379
; CURRENT APPLICATION NUMBER: US/10/055,364
; CURRENT FILING DATE: 2002-01-23
; PRIOR FILING DATE: 2000-07-07
; PRIOR FILING DATE: 2000-07-07
; PRIOR FILING DATE: 1999-07-08
; PRIOR FILING DATE: 1999-07-08
; PRIOR FILING DATE: 1999-12-02
; NUMBER OF SEQ ID NOS: 55
; SOFTWARE: Patent in version 3.0
; SEQ ID NO 48
; LENGTH: 359
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Suid herpesvirus 1 - bases 491-850
US-10-055-364-48

Query Match
Best Local Similarity 9.7%; Score 442; DB 14; Length 359;
Matches 108; Conservative 83; Mismatches 152; Indels 24; Gaps 8;

; 441 SRRSRQ--AESTDPMWNTGNGAGGESSNSITVAQVQYAYNLIRINNIILEDLSK 498
; 6 ARRARESPGAGTPEP---PAVNGTGHRLITTSAGFAERLQFTYDHIQAHVNDMLGRTAA 62
; 499 AWCREQRAALVNNELSKINPTSVMSIYNRPVSARKIGDVISVNCNIWDQTSVSLHKS 558
; 63 ANCELQNKORTLWSENRLNPAVATAALQQRVSRMLGDVNAISRCVEV-RGVYVQNS 121
; 559 LRLLSASDEKCFSPRPVTFKFNMDSTIYKQGLGVNNEILLTTLTETCOENTYFFQAKT 618
; 122 MR-VPGERGTCSYRPLVTFEH-NGTVIEQGLDGNELLISRDLEPCGTGNHRRYFKLGS 179
; 619 DMWIKYKNEHLKTVPLSSITLDTFTALNFTLLENVDFKVIYLRDEKRLSNWFDIETM 678
; 180 GYVYVEDYNYVRMVEYPE--TISTRTVLTNLTLEDREFLEPLFVYTRTELADTGLLDYSEI 237
; 679 FREYNYVA-----QRVSLGRKDLDDLSTNRNQFVDAFGLSMDLGGVQTVNNAVSQA 732
; 238 QRRNQHLAKFYDIDRVKVDNV
```



```
; SEQ ID NO 2
; LENGTH: 3907
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-171-311-2

Query Match      2.8%; Score 126.5; DB 15; Length 3907;
Best Local Similarity 18.7%; Pred. No. 1.5;
Matches 110; Conservative 101; Mismatches 196; Indels 181; Gaps 26;

QY 373 KEFTTYHCLNEIEKAEIEKKYAKVNSTHSGYDLYKPKTDGGLYLVWQPLIQNRLLDAAK 432
Db 605 KEKNAVLDMAESQEAELERLTQLLSHEE--ELSKLKED-----LEIE 647
QY 433 NKLNNETYRRSRRAQESTTDPMMEMTGNGAGGEY-----SSENSITVAQVQAYDNL 485
Db 648 HRIN-----IEKLDNLGIHYKQIDGLQNMESQKIETMQEKNL 688
QY 486 RIRINNILEDLSKAWCREQ-----HRAALVNNELSK-----INPT 520
Db 689 ITKQNLILEISKDLQSLVNSKSEEMTLQINELQKEIEILRQEEKEKGTLEQVQEL 748
QY 521 SVMSMYINRPVSAKRIGDV-----ISVSNCLVVDQ-----TSVSLHKSRLILS 563
Db 749 QUKTELLEKQMKKE--NDLQEKFAQLEAENSILKDEKKTLEDMLKIHTFPVQOERLIFLD 807
QY 564 ASDEKCFSPVPVTFK----FMNDSTIYKGO--LGVNNEILLTTTLYLETQENTYVYFQ--- 615
Db 808 SIKSK--SKDSWEKEIEILIEENEDLKQCCIQLNEEIEKORNTFSPAKEKFEVNYQELQ 865
QY 616 -----AKTDMYIYKNYEHK-----TVPL--SSITLTD--TFIA 645
Db 866 EYVACLLKVKDDLEDKQKQLEYSKALKALNEELHLQRIINPTTVKMKSSVDFDEKTFVA 925
QY 646 LNFTLLENVDKVIEL-----YTRDEK-----RLSNVFD--IETMPREYNYAQRVSGLRK 694
Db 926 ETLGEMGEVVEKDTTLMKLEKLEVTREKLELSQRLSDLSQKQKGEISFLNEEVKSLQ 985
QY 695 DL-----LDLSTNRNQFVDAFGSLMDLDGAGQTVVNAVSGVATLFSISVITGFINF 745
Db 986 EKEQVSLRCRELEIIINHNR-----AENVQSCDTQVSSLLDGVVMTSGAGSGVSK 1037
QY 746 IKNPFGGMLMIIV-----VIG-----VLFAIYFLTKTKIYETAPIK--MIYP 786
Db 1038 VNKSFGESKIMVEDKVSFENMTVGEESKQQLILDHLPSTVKESLRATQPSENDKLOK 1097

787 EIDKLKEREKGEIAPISEEELERIVLAM-----HIHQONSHMETKTRK 830
1098 ELNVLKSEQNDLRL-----QMEAQRICLSLVYSTHVDQVREYMEKDK 1141

RESULT 26
US-10-171-311-6
; Sequence 6, Application US/10171311
; Publication No. US20030087270A1
; GENERAL INFORMATION:
; APPLICANT: Schlegel, Robert
; APPLICANT: Chen, Yan
; APPLICANT: Zhao, Xumei
; APPLICANT: Monahan, John
; APPLICANT: Kamatkar, Shubhangi
; APPLICANT: Glatt, Karen
; APPLICANT: Gannavarapu, Manjula
; APPLICANT: Hoerish, Sebastian
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY
; FILE REFERENCE: MRI-035
; CURRENT APPLICATION NUMBER: US/10171,311
; PRIOR FILING DATE: 2002-06-12
; PRIOR APPLICATION NUMBER: US 60/298,159
; PRIOR FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/298,155
```

```
; CURRENT APPLICATION NUMBER: US/10/171.311
; CURRENT FILING DATE: 2002-06-12
; PRIOR APPLICATION NUMBER: US 60/298,159
; PRIOR FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/298,155
; PRIOR FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/335,936
; PRIOR FILING DATE: 2001-11-14
; NUMBER OF SEQ ID NOS: 238
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 3925
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-171-311-6

Query Match      2.8%; Score 126.5; DB 15; Length 3925;
Best Local Similarity 18.7%; Pred. No. 1.5;
Matches 110; Conservative 101; Mismatches 196; Indels 181; Gaps 26;

QY 373 KEFNTTVHCLNEETKAEIEKKYAKVNSTHSGYGLKYFKTDGGLYLVWQPLIQNRLLDK 432
Db 605 KEKAVLDMAESQEAELERLURTLQFSHEE--ELSKLKD-----LEIE 647
QY 433 NKLANNYSRRSRQAESTTDPMMEMTNGAGGEY-----SSENSITVAQVAYDNL 485
Db 648 HRIN-----IEKLKDNLGIHYKQIQIDGLQNMESQKIETMQFEKDNL 688
QY 486 RIRINNILEDLSKAWCREQ-----HRAALVWNELSK-----INPT 520
Db 689 ITKONQILFLEISKLDQOOSLVNSKSEMTLQINELQKEIEILRQEBKEKGTLEQVEQL 748
QY 521 SVMSMIVNRPVSXKRGIV-----ISVSNICIVDO-----TSVSLHKSRLLS 563
Db 749 QKTELLEKQWKEKE--NDLQEKFAQEAENSILKDEKKTLEDMKIHTPVSQERLIFLD 807
QY 564 ASDEKCFSRPVPVTFK-----FMNDSTIYKGO--LGVNNEILLTITYLETCQENTYFQ--- 615
Db 808 SIKSK--SKDSVWEKEIEILTEENEDLKQOCIQLNEEIEKORNTFSAEKNFEVNYQELQ 865
QY 516 -----AKTDMYIKNYEHLK-----TVPL--SSITLID--TFIA 645
Db 866 EYACLLKVKDDLEDSKNQOEYKSKLKALNEELHLQRIINPTTVKMKSSVFDDEKTFVA 925
QY 646 LNFTLLENVDPKVIEL-----YTRDEK-----RLSNVFD--IETWFRVNYAQRVSLRK 694
Db 926 ETLGEVVEKDTTLEMEKLEVTREKLELSQRULDSLEQKQKHGEISFLNEEVKSLQ 985
QY 695 DL-----LDLSTNRNQFVDAFGSLMDDLGAVGQTVVNAVSGVATLFSSIVTGFINF 745
Db 986 EKEQVSLRCRELEIIINHNR-----AENVQSCDTQVSSLDGVVVTMSRGAEGSVSK 1037
QY 746 IKNPFPGMLTIIV-----VIG-----VLFAYIFLTKTKIYETAPIK--MIYP 786
Db 1038 VNKSGFEESKIMVEDKVSFENMTVGESKOBQLILDHLPSTVBESSRATQPSENDKLQK 1097
QY 787 BIDKLKEREKSEAPISEELERIVLAM-----HIHQONSHMETKTK 830
Db 1098 ELNVLKSEQNDLRL-----QMEAQRICLSLVSTHVDQVREYMEKDK 1141
```

RESULT 27

```
US-09-815-242-12141
; Sequence 12141 Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
```

```
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12141
; LENGTH: 1009
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-09-815-242-12141
```

```
Query Match      2.7%; Score 125; DB 9; Length 1009;
Best Local Similarity 17.3%; Pred. No. 0.24;
Matches 114; Conservative 112; Mismatches 231; Indels 202; Gaps 30;

QY 267 EMSPPWSGEDDHENKHEKWPFSVINNVK--VDYQNRGTVPGLKTRIFLDREYTLISWE 325
Db 459 ELDKLGSKQOLENQEENEEK--TILKNLEIKLDYDKSKLDLN-----DKESFISEIK 509
QY 326 KHLKNMSYPLTLWKAIFYNGIOFEGHSGSYHFVANDITASFTTSKEDMKENFTTYHCL-- 382
Db 510 SAVKIGDQCPIC-----GNEIQDILG--HHIDFD--SIKRONKEIKEIANIHWMESN 557
QY 383 ----NEEIKAEIEKKYAKVNSTHSGYGLKYFKTDGGLYLVWQPLIQNRLLDKAKNLNE 438
Db 558 IAVHNSKIKF-VNEKISNIN-----IKTQSDLSL--EVLNKRLLLENALNNQ 602
QY 439 ----TYSRRSRQAESTT-----DPMEMTNGAG-----GEYSSENSITVA 476
Db 603 RLNKFIEQMKKEKONLTQIHNKQLRLNKNSELSKICRNLITFETLTKYNNITNFEVD 662
QY 477 QVOYAYD-NLRIRINNILED-----LSKAWCREQRAALVWNELSKINPTSVMSMIVNRPV 531
Db 663 YKXIQDVNQHQHSHSQIEDKLIQLSQRKLIQONNLNHYENQLE-----TYNNDL 712
QY 532 SAKRIGDVISVSNICIVVDQTSV-----IIAWRGEQBELEQKRDYKRYHEFEMEIALESLT 556
Db 713 ELNEQSIEMSKRLNLTDDNDINEIIAWRGEQBELEQKRDYKRYHEFEMEIALESLT 772
QY 557 KSLRLLLSASDEKCFSPRPVTFKPMNDSTIYKQOLGV-----NN-----EILL 598
Db 773 KOKELLDSD-----KLKDDYELKKGKMTLIDEYSAAVHYQCQNNIKKTQSIVS 820
QY 599 TTYLETQCOENTYFYQAKTDMYIKVNYEHLKTVPLSSITTLDTFTALNFTLLENVDFKV 658
Db 821 HYNLNQELKDOQEIPQL-AEILSGKNNKL-----TLE-----NFVLIYYID-- 862
QY 659 IELYTRDEKRLSNVFDIETMFREYNYAQRVSGLRKXDLDLSTNRNQFVDVAFGLSLMDLG 718
Db 863 -QIIAQANLRLATMSDNRYQLIRREAVSHGLSGLEIDVDFLHNSKSRH----- 909
QY 719 AVGQTVNAVSGVATLFSS--IVTGFINKPFGGMLMIIVVIGVLFALYFLTKTKIY 776
Db 910 -----ISSLGGETFQSSALALGLSEIVQOOSGGISLESIFIDEGFG---TLDQETL 959
QY 777 ETAPIKMIYPEIDKLKEREGKSE----IAPISEEUEERIVLAMHIHQONSHMETKTKRD 831
```

Db 960 ETV-----LDTLLNLKSTGRMVGLIISHVSELK-NRIPLVLEVKSDQYQSSTFRKN 1009

RESULT 28

US-09-815-242-5251

; Sequence 5251, Application US/09815242

; Patent No. US20020061569A1

; GENERAL INFORMATION:

; APPLICANT: Haselbeck, Robert

; APPLICANT: Ohlsen, Kari L.

; APPLICANT: Zyskind, Judith W.

; APPLICANT: Wall, Daniel

; APPLICANT: Trawick, John D.

; APPLICANT: Cart, Grant J.

; APPLICANT: Yamamoto, Robert T.

; APPLICANT: Xu, H. Howard

; TITLE OF INVENTION: Identification of Essential Genes in

; FILE REFERENCE: ELITRA.011A

; CURRENT APPLICATION NUMBER: US/09/815,242

; PRIOR FILING DATE: 2001-03-21

; PRIOR APPLICATION NUMBER: 60/191,078

; PRIOR FILING DATE: 2000-03-21

; PRIOR APPLICATION NUMBER: 60/206,848

; PRIOR FILING DATE: 2000-05-23

; PRIOR APPLICATION NUMBER: 60/207,727

; PRIOR FILING DATE: 2000-05-26

; PRIOR APPLICATION NUMBER: 60/242,578

; PRIOR FILING DATE: 2000-10-23

; PRIOR APPLICATION NUMBER: 60/253,625

; PRIOR FILING DATE: 2000-11-27

; PRIOR APPLICATION NUMBER: 60/257,931

; PRIOR FILING DATE: 2000-12-22

; PRIOR APPLICATION NUMBER: 60/269,308

; PRIOR FILING DATE: 2001-02-16

; NUMBER OF SEQ ID NOS: 14110

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 5251

; LENGTH: 996

; TYPE: PRT

; ORGANISM: Staphylococcus aureus

US-09-815-242-5251

Query Match 2.7%; Score 124; DB 9; Length 996;

Best Local Similarity 17.4%; Pred. No. 0.28;

Matches 109; Conservative 106; Mismatches 216; Indels 196; Gaps 28;

QY 267 EMSFWSGDDHKNMKHFKPVSVINNVKV-VDYQNRGTVPGLKTRIFLDREEVTLSE 325

Db 459 ELDKLLGSQLENGINEEK---TILKLEIKLDYDKSKLDN-----DKESFISEIK 509

QY 326 KHLKMSVCPLTLWKAIFYNGIQTEHSGSVHFVANDITASFTTSKEDMKFNTTYHCL--- 382

Db 510 SAVKIGDQPCIC-----GNEIQDLG--HHIDFD---SIAKRQNEIKEIEANIHTMESN 557

QY 383 ----NEEIKAEIKKYAKVNSTHSGYDLYKFTGGYLVWQPLIQNRLLDAKKNLNE 438

Db 558 IAVHNSIKF-VNEKISIN-----IKTQSDLSL-----EVLNKRLLLENALNNQ 602

QY 439 ----TYSRRSRQABSTT-----DPNMTGTGNAG-----GEYSSENSITVA 476

Db 603 RDLNFKIEQMKSEKNTLIQHNKQLRLNKNSELIKICNLITEFETLSKYNNTNFEVD 662

QY 477 QVOYAYD-NLRIRINILED-----LSKAWCREQHRAALVWNELSKINPTSVMSMYNRPV 531

Db 663 YKKYIQDVNQHQHSHSNOIEDKLIQSORKLIENNNLHYENQLE-----TYNNDL 712

QY 532 SAKRGDVLVSNVICVVDTSV-----SLH 556

Db 713 ELNEQSIEMEMRLNLTDDNDINEIIAWHGEQEELEQKRDYKRYHBFEMEIALESIT 772

QY 557 KSLRLLSASDEKCFRPPVTFKFMNDSTIYKQLGV-----NN-----EILL 598

Db 773 KDRELLDSD-----KLKDDYELKKGKNTLIDEVSAVHYQCONNIKKTQSIVS 820

QY 599 TTTYLETCEQNTYFYFOAKTDMYIYKNYEHKLTVPVLSSTITLDTFIALNFTLLENVDFKV 658

Db 821 HINYLQELKQDQEIFOL-ABILSGKNKNL-----TLE-----NFVLIYLD--- 862

QY 659 IELVTRDEKRLSNVFDIETMFREYNYAQRVSGRLKDLDLSTNRNOFVDFAFGSLMDDL 718

Db 863 -QIIAQANLRLATMSDNRVQLIRREAVSHGSLGLEIDVFDLHNSKSH- 909

QY 719 AVGQTVNAVSGVATLFSS--IVTGFINFINKPFGGMLMIIVIGVLFAYIYFLTKT--- 773

Db 910 -----ISSLSGGETFOSSLALALGLSEIVQQQSGGISLESIFIDEFGT--LQOETLET 961

QY 774 -----KIYETAPIKMIYPEIDKLER 794

Db 962 VLDTLNLKSTGRMVGIISHVSELKNR 988

RESULT 29

US-10-242-056-59

; Sequence 59, Application US/10242056

; Publication No. US20030113323A1

; GENERAL INFORMATION:

; APPLICANT: Ensign, Jerald C

; APPLICANT: Bowen, David J

; APPLICANT: Petell, James

; APPLICANT: Fatio, Raymond

; APPLICANT: Schoonover, Sue

; APPLICANT: French-Constant, Richard

; APPLICANT: Orr, Gregory L

; APPLICANT: Merlo, Donald J

; APPLICANT: Roberts, Jean L

; APPLICANT: Rocheleau, Thomas A

; TITLE OF INVENTION: Insecticidal Protein Toxins from

; NUMBER OF SEQUENCES: 88

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: DowElanco

; STREET: 9330 Zionsville Road

; CITY: Indianapolis

; STATE: IN

; COUNTRY: US

; ZIP: 46268

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; FILING DATE: US/10/242,056

; CLASSIFICATION:

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/063,615

; FILING DATE: 18-MAY-1993

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/395,497

; FILING DATE: 28-FEB-1995

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 60/007,255

; FILING DATE: 06-NOV-1995

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/608,423

; FILING DATE: 28-FEB-1996

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/705,484

; FILING DATE: 28-AUG-1996

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/743,699

; FILING DATE: 06-NOV-1996

; ATTORNEY/AGENT INFORMATION:

NAME: Borucki, Andrea T.
REGISTRATION NUMBER: 33651
REFERENCE/DOCKET NUMBER: 50301E
TELECOMMUNICATION INFORMATION:
TELEPHONE: 317-337-4846
TELEFAX: 317-337-4847
INFORMATION FOR SEQ ID NO: 59:
SEQUENCE CHARACTERISTICS:
LENGTH: 1565 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
us-10-242-056-59

Query Match 2.7%; Score 124; DB 15; Length 1565;
Best Local Similarity 17.9%; Pred. No. 0.58; 314; Indels 386; Gaps 46;
Matches 179; Conservative 119; Mismatches 314; Indels 386; Gaps 46;
36 PNTATWSTESPLTGHYTHSSHGGERGN--NENR-----DSEEQKN 75
3289 PNTTTLMTQQ-----AGESSETQSSLLIDESTTLRQVNLATTFDSIDPTEETDSN 341
76 IYG-----SPSTFPYRVCSAGVGDFRPQTDHVCDFDASDMVHSE 115
342 PYGRMLGLGVFRQFEGDCANKRKNPVVYGYLYCDSAFNRHVLK-----PLSKNFLFS- 393
116 GILLIYK-----QNIIPFMFRVKY--RKVVT-----TSTVYNG 147
394 ----TYRDETQGNLSQFAVYDKKYITKVVTGATEDPENTGWVSKVDDLKQGTGAYVY 449
148 IYDSIT-----NQHTEKYSIEPWETEKMDTIYQCFNSLRNLNTGGLLTVDVR 195
450 IDQDGLLHIQTTNGDFINRHTFGYN-----DLVYDSKSGYGFWSGEGFVLDY 500
196 DDINNTVLPQVGVGTPDVVKRYGQSP-----EULYEPGFMGYSRRRTTNCCLMDMFA 249
501 HDGNYTFHNAIINYPSPGYGGVPGTWALEQINEGW-----AIAPLLDTL- 549
250 RSNPPDFFTATGDTVMSPFWSGEDDHENKHEKPFVSVINNYKVVDYQNRGTVPLG 309
550 -----HTVTVKGSYTA-----WEGETPTGYNLY-----IPGTVLLDWFQKINFAIG 591
310 KTR---IFLDREYTLTWEKHLKMSYCPLTLKAFYNGIQTEHSGSYHFVANDITASFT 366
592 LNKLESVFTSPDWFTLA--TIKFS--KIADNRKFYQEIINAEATADGRNLFKRYSTOTFG 646
367 TSKEDMKFNTYHCLNEEKAEIEKKYAKV-----NSTHKKYGDLLKYPKTDGGLY 417
647 LTSG--ATYSTTYTLSEADFSTDPDKNYLQVCLNVVWVDHYDRSPGKKGAYSWSVKWFNVY 704
418 LVMO-----PLIQNRL-----LDKKNLANETYSRRSRQAE---- 449
705 VALQDSKAPDAIPRLVSKYSDSKGLVLOYLFWTSSLPKATRLNT-TPVRTLIEKANIGLD 763
450 -----STTDPFME-----MTGNGAGGEY-----SSENSTIVA 476
764 SLDDYTLQADPSLEADLVTDGKSEPMDFNGSNGLYFWELFFHLPFLVATRFANEQQFSPA 823
477 Q--VOYAYDNLIRINNILEDLSKAWCREQHRALVWN-----ELSK-----INPT 520
824 QKSLHYIFD-----PAMKNKPHNAPAYVWVRPLVEGNSDLRHLDDSDIDPD 869
521 SVMSMIYNNRPSAKRIGDIVSVNCIV-----VDQTSVSLHSLRLLSASDE 567
870 ---TOAVAHPIYOKAVFIAYVSNLIAQGMWYRQLTRDGLTQARVYVNYNLAELLG---- 922
568 KCFSRPPVTFKFM-----NDSTIYKQGLGV-----NNEILL-----TTTYLETQCE 608
923 ---PRPDVSLSSIWTPQTLDTLAAGQAKVLRDFEHQLANSDTALPALPGRNVSYLKLDN 979
609 NTEYFQAKTDMYIKNYEHLK-----
980 G---YFNEPLNLVLMLSHWDLTDLARLYNLRNLHTVDGKPLSLPLLYAAPVDPVALLAORAQS 1036

QY 631 -----TVP-----LSSITTLDTFIALNFTLENVDFKVIYLRDEK 667
DB 1037 GILTVNGVSGAMLTVPYPRFSAMLPRAYSAVGTLTSGQNLLSLLER-----SRAQOEEL 1091
QY 668 RLNSNVFEDIETMFREYNYA-----QRVSGLRKDLDDLSTNR-----NOFVDAFGLS-MDDL 718
DB 1092 AQOQLDMSS-----YAITLQQQALDGLAADRLALLASQATQQRHHDHYTYLYQNNIS 1144
QY 719 AVQOTVNAVSGVATLFSS-----IVTGFINFIKNPF 751
DB 1145 SAEQLVMDTQTSQAQLSSSTGVQTASGALKVIPNIFG 1182

RESULT 30
US-09-815-242-12327
; Sequence 12327, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12327
; LENGTH: 792
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-09-815-242-12327

Query Match 2.7%; Score 123; DB 9; Length 792;
Best Local Similarity 18.2%; Pred. No. 0.24;
Matches 108; Conservative 93; Mismatches 250; Indels 144; Gaps 17;
QY 346 IOTEGSHGVHFVANDITASFTTSKEDMKFNTYHCLNEEKAEIEKKYAKVNSTHSHKYG 405
DB 183 INHQHRDVAKVALENIKAWFGSFNEKMSERNQEKQLKREE-KARLKEEQKARQEQPIK 241
QY 406 DLK-----YFKTDGGLYLVWQPLQNLRLLDKKNLANETYSRRSRQAEETT 452
DB 242 DVSDFTVEPOERDIPITYGHTENESKSQSPSRKRVFDAENSSNNIVNHHQADQSQSLT- 300
QY 453 DPMWEMTNGAGGEYSSENSITVAQVQYAYDNLIRINNILEDLSKAWCREQHRALVWN 512
DB 301 ----EQTHNSVSENTEIEAGEVTNVSVYVPLTL-LNQPAKQKATSKAEVQKQGVLEN 355
QY 513 ELS-----KINPTSVMSMIYNNRPSAKRIGDIVSVNCIVVDQTSVSLHSLRL 561
DB 356 TLKDFGVNAKVTQIKIGPAVTQYEI--QPAQGVKVKSVIYNLHNDIAL-----A 401

APPLICANT: Institut National de la Recherche Agronomique
; TITLE OF INVENTION: Resistance Mechanisms to IC Type R/M
; FILE REFERENCE: Bacteriophages of Lactic Acid Bacteria
; CURRENT APPLICATION NUMBER: US/09/486,734A
; CURRENT FILING DATE: 2000-05-03
; PRIOR APPLICATION NUMBER: PCT/FR98/01873
; PRIOR FILING DATE: 1998-09-01
; PRIOR APPLICATION NUMBER: FR 97/10885
; PRIOR FILING DATE: 1997-09-02
; NUMBER OF SEQ ID NOS: 40
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 1025
; TYPE: PRT
; ORGANISM: Lactococcus lactis
US-09-486-734A-4

Query Match 2.6%; Score 117.5; DB 10; Length 1025;
Best Local Similarity 17.5%; Pred. No. 1.1;
Matches 173; Conservative 163; Mismatches 314; Indels 339; Gaps 49;

QY 61 RGNN-----ENRDEE-----QNKNIYGSPTPPYRVCSASGVGVFRFQTDHVPD 107
DB 83 RCGNGVAQVLLEREDGKVTLEAFRNKDISGTS--YEV-----VHQVVPD 127
QY 108 ASDMVHSEGILLI-----YKQNIIPFMFRVRKYRVKVTSTTVNGIYSOS- 152
DB 128 SRVRGDSVSLINGLPIITHIELQESAKDGMQAYIQIRAE-----DGFPGIYATQ 183
QY 153 ---ITNQ-HTFY-----KSIPEWETEKMDTIYQCFN-----SLRLNTGNN 188
DB 184 INVIPNKVTRFYARPSDEATAEAYARMKFLFNWRTEDNQTDJLDFTRTVLRIPDAHE 243
QY 189 LUTYV-----DRDDINMTVLPD-----VDGTPDVVKYGSQPELYLEPGFW- 235
DB 244 LISQVTLVDDPKNPKFLMALRPYQIHAIKIRKRAAQH-----EGGFIWHATGSGK 295
QY 236 R-----RTTVNCELMDMFARSPPEDFVTA--TGDVMSFPWSGE 275
DB 296 TITSFVATKLLAQAIGVDRVTVMVDRDLDQAQDEFTKFASEYHTGQTGNSVANTLI 355
QY 276 DHENKMKHPFVSVINNYKV-----DYQNRGTVPGLKTR-----IPL 315
DB 356 VGKQKQKQLARNLLSKNNNTILVTIQKLSAAMSAQOESEKSGNQPEKLQEHIVFI 415
QY 316 DREBYTLSEKHLKQM-SYCPITLM-----KAFYNGIQTEHSGSVHFVAND-----ITASF 365
DB 416 VDEAHRAVSDEEMKRIKKILPNSTWFGLTGTPIFEANKKQENGTEARTTSQQYGPLLSY 475
QY 366 TTSKEDMKE-----FNTTYHCL-----NE 384
DB 476 TT-KNAMDGAVLGFQVEYHSLISEDBLEVIYQLNKGKLPDALQOEELLPAEYKDE 534
QY 385 EIKAEIEKKY-----AKVNSTHS-----KYGDKYFKTDGGLYLW 420
DB 535 HIRTMLOKIFNRSVVVKFKVNGPPTMSAILTTHSIAQAQHIYRLKEMKONG-----588
QY 421 QPLIQNRLLDAKNKLNNETYSR-----RERRQAESTDDPMEMTNGAGGEYSSENS 472
DB 589 -TLLNGRQFDERHRLDKDPFRAITFTSTNPDRLEKNEQDDELVEIM-----KEYAKQPD 642
QY 473 ITVAQVOYADNLRIRINILEDLSKAWCRBQHRALVNE-----LSKINTSVMS 524
DB 643 ASYQDEKLYN-----QINIKRLARK--EKQYQSDGQWLDFVIVVDRLLTGFDSPAIT 694
QY 525 MIYNRPVSAKRIGDVISVNCIWDOTS---VSLHK-----SLRLLSASDEKCFPS 571
DB 695 LYIDRENNYQKLLQAFSRNRYITGKDSGLIYFSPKPTMKENVQNTFLFNEHQNFQ 754
QY 572 RPP-----VTFKPMNDSTIYK---GOLGVNNEILLTITYLETQENETYYFOAKTDMYI 623
DB 755 LIPREYEEVKFEICSTLYKQSEADLSDNPHDLKTMIAQVSAQYKLGKSYKA-----F 808

RESULT 35

US-10-087-464-10
; Sequence 10, Application US/10087464
; Publication No. US20030059436A1
; GENERAL INFORMATION:
; APPLICANT: Chishti, Athar
; APPLICANT: Oh, Steven
; APPLICANT: Liu, David
; APPLICANT: Goel, Vikas
; APPLICANT: Li, Xuerong
; TITLE OF INVENTION: Band 3 Antigenic Peptides, Malaria Polypeptides and Uses Thereof
; FILE REFERENCE: SI237/7019
; CURRENT APPLICATION NUMBER: US/10/087,464
; CURRENT FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: US 06/272,930
; PRIOR FILING DATE: 2001-03-02
; NUMBER OF SEQ ID NOS: 59
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 10
; LENGTH: 1639
; TYPE: PRT
; ORGANISM: Plasmodium falciparum
US-10-087-464-10

Query Match 2.6%; Score 117; DB 15; Length 1639;
Best Local Similarity 18.6%; Pred. No. 2.5;
Matches 148; Conservative 122; Mismatches 254; Indels 270; Gaps 39;

QY 55 DSSHGERGNN-----ENRDSSEQNKNIYGSPTPPYRVCSASGVGVFRFQTDHVPDAS 109
DB 912 DTSHSTNLNLSLKLFPENITSLGKKNKIY-----939
QY 110 DMVHSEGILLIYKQNIIPFMFRVRKYRVKVTSTTVNGIYSDSITNQHTFYKSIPEWETE 169
DB 940 -----QELIGQKSSSENFYEKILKDSDTF-----YNESFTN----FVKS-----973
QY 170 KMDTIYQCFNSLRNTGNNLTYYVDRDINMTVLPQVDPVDPV-KRYGSQPELYLBP 228
DB 974 KADDI-----NSLDESRRKKL-----EEDINK--LKKTLQLSFDLYNKYKLERLFPKK 1022
QY 229 WFWGYSR---BRTTVNCELMD--MFARSNPP-----PDFF-----VTATGDTVMSP 270
DB 1023 KTVGKYKMQIKKLTLLKEQLESKLNSLNNPKHVLQNFVSFFNKKEAEIAETENTLENTK 1082
QY 271 FWSGEDDHENKMKHPFVSVINNYK-VVDYQNRGTVPGLKTRIFLDREBYTLSEKHLK 329
DB 1083 I-----LLKHYKGLVRYNGESSPL-----KTLSEESIQT 1112
QY 330 NWSYCPILTKAFYNGIQTEHSGSVHFVANDITASFTTSKEDMKEFNTTYHCLNEEIKAE 389
DB 1113 EDNVASLENFVKLSK-----LEGKLNKLNLEKKKLSYLSGLHLLAELKEV 1160
QY 390 IEKKYAKVNSTHSKYGDL-----KYFKTDPGGLYLVNQPLIQNRLLDAKNKLNNETYSR 442
DB 1161 IKKNYTCGNSPSENNTDNNNALESYKFKLPEG-----TDVATVVSSESGSDTLEQ 1209


```
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; TITLE OF INVENTION: Prokaryotes
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11522
; LENGTH: 1167
; TYPE: PR1
; ORGANISM: Helicobacter pylori
US-09-815-242-11522

Query Match      2.5%; Score 115.5; DB 9; Length 1167;
Best Local Similarity 17.7%; Pred. No. 1.9;
Matches 151; Conservative 118; Mismatches 267; Indels 319; Gaps 36;

QY 51 YGTHDSSHGRCNNEN-----RDSEQKNYIG-SPSTFPYRVCASGVDPFRFQ 100
DB 270 YNMFYKLGDRNYENFAKKNIAARTLNERKULFGKNPEFDNFLTSLR--GNIHOSI 327
QY 101 TDHVCPCDASDMVHSEGIILYKQNIIPFMRVRYKRVKVVTTSTVYNGIYSDSITN- 155
DB 328 KEE---EALDWTIISH-----LITKPIEDAIFGDNKINPIAKA 361
QY 156 -----QHTFYKSTIEPNET 168
DB 362 LDKMWLKLSDLGETKDKLKNLYESVTEARAKSQSKQBELIKNLYNTFFKEAFKQOS 421
QY 169 EKMDTIY---OCFNSLRNTGGLTYVDRD--DINMTVELQPDVGVTPDVKRYCSQBEL 223
DB 422 EKLGIVYTPIEVVDFILRATNGILKKHFTNFDNQSITIP-DPFWGTGSGFIARLLSKEND 480
QY 224 YLEPGWFGSVRRRTTVNCELMDMFARSNPPDFVFVATGDTVEMSPWSEGDHDKHMKH 283
DB 481 FIS-----DEALKEKFQKGLFAPDIVLLS----- 504
QY 284 EKPWFVSVINNVYVDYQNRGTVPGLKTRIFLDREYTLSEKHLKNMSYCPCLTLWKAFY 343
DB 505 ---YVIALINITQAA--QNR-----DSSLKKNFKNIALT----- 532
QY 344 NGIQTEHSGSYHFVAND--ITASPTTSKEDMKFNTYTHCLNEEIKAEIKYKVNSTH 401
DB 533 -----DSDLYLEEKNDKGVIPGPFYLPEDLKE-----NKEIKTMMKQNRVITGN 578
QY 402 SKYGDLYKFKYTDGGLYLVWQPLIQNRLLDAKNKLNNETYSRRSRQ-AESTTDPMMWTG 460
DB 579 PYPYSGAKSENNDNQNL-SHPKLEKRVY-----EYKGNSTAKQNKSTRTDLI----- 625
QY 461 NGAGEYSENSITVAQVOYVDNLRIRINNILEDLSKAWCREQHRRAALVWNELSKINPT 520
DB 626 -----HSIRWA-----SLLKDK-----K 638
QY 521 SVMSMIYNRPVSAKIGDVISVSNICIVVDQT---SVSLHKSRLLSASDEK-----CF 570
DB 639 GVLGFWNVGNSFIDSKSAD--GPRKCAQDFSHLYALNLRGNARTSGEERKQGGGIDFSG 696
```

```
571 SRPPVTFKFMNDSTIYKQGLGVNNEILLTTTYLETCEQNTYFYFOAKTDMYIYKNYEHKL 630
697 SRVTVAIIF-----FVKDKDAPNHTIF-----YVE-----VEDYLKREAKNLLANFENDL 742
631 TVPLSSITTLDTFTALNFTLLENVDFKVIELYTTDEKR--LSNVFPIETMREYNYAQR 688
743 SVPFKEITPNDKGDWIN---QRNDDFEKLIPLKRDKKSKIFNAIFDLNS----- 788
689 VSLGRKDLLDLSTNRNOFVDAFGSLMDLDGAVGQTVVNAVSVATLRFSSIVTGFINFTKN 748
789 -NGVK-----TSRDPWVYNFSQ-----KTLMSQVQNCIDTVDNADLKRNERPRE 831
749 PF-----GGMLMIIVVIGVLFAYFLTKTKIYETAPIKMIYPEIDKLKEREGKSEIA 801
832 AFQRTAKDKGKKSADRYKHLNDREITTDKTIKAWT-----DGLANKLIKNEUL 881
802 PISEEELERIVIAWH 816
882 P--ESGMERVELALY 894

RESULT 38
US-09-732-180-7
; Sequence 7, Application US/09732180
; Patent No. US20020137702A1
; GENERAL INFORMATION:
; APPLICANT: Arkowitz, Robert A
; APPLICANT: Neirn, Peter MA
; TITLE OF INVENTION: Yeast receptor
; FILE REFERENCE: DY0025.001AUS
; CURRENT APPLICATION NUMBER: US/09/732,180
; CURRENT FILING DATE: 2000-12-07
; PRIOR APPLICATION NUMBER: US 60/169,699
; PRIOR FILING DATE: 1999-12-07
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 7
; LENGTH: 813
; TYPE: PR1
; ORGANISM: Candida albicans
US-09-732-180-7

Query Match      2.5%; Score 115; DB 10; Length 813;
Best Local Similarity 19.2%; Pred. No. 1.2;
Matches 150; Conservative 107; Mismatches 267; Indels 256; Gaps 40;

QY 75 NIYGSPTFPYRVCASGVGVDFRQTDHVCPCDASDMVHSEGIILYKQNIIPFMRVRYK 134
DB 148 NISVDPATHLWKLFQ---GAPFCVLINHLIPDSQIPVWSSDDLRLCKSVYDFLIAVKT 204
QY 135 Y-----RKVVTTSTVYNGIYSDSITNQHFTFYKSIETEKMDTIYQCFNSLRNTGGL 189
DB 205 QLNFDDENMFTISNVFSDNAODLI-----KIIDVINKL-----L 238
QY 190 LTYVDR-----DDINMTVFLQPDV-----GVTPDVKRYGSOPELYLEPGWFGSYR 235
DB 239 AEYSDASDSGGEDVNMV--QITDERSKVFREIIEETERYKIVODLELMCK-----Y 288
QY 236 RRTTVNCELMD-----MFARSNPPDF-----FVTATGDT 265
DB 289 RDLIEAENLSSEQIHLLFPNLNEIIDFORFLNGLECNINVPTRYGRIGSVFIHAS--- 345
QY 266 VEMSPFWSGEDDHENKHEKFPWFSVINNVYVDYQNRGTVPGLKTRIFLD-----REY 321
DB 346 --LGPFNAYE-----PW---TIGQLTAIDLINKEAANKSSSLDPGFELOS 390
QY 322 LSWEKHLKNMSYCPCLTLWKAFYN---GIQTEHSGSYHFVANDITASFTTSKEDMKFNT 377
DB 391 L---KPIQRICKYPLLLKELIKTSPEYSKODPHGSSSSTSFNELLVAKTAKMELANQN- 446
QY 378 TYHCLNEEIKAEIKYKVNSTHSGYDLKYFKTDGGLYLVWQPLIQNRLLDAKNKLN 437
DB 447 -----EAQRAENIEHLEKLK---ERVGNVRGNFLDAQGELLFHG--QVGVKDAENEKEY 496
```

Qy 438 ETY-----SRRSROAESTTDPMMEMTNGAGGEYSSSENSITVAQVQYAYD 483
Db 497 VAYLFEKIVFFFTTIDTKSDQKEKKS-----KFSTRKRSTSSNLSSSTT 542
Qy 484 NLRIRINILEDLSKAWCREQHRAALVWNELS-KINPTSVMSMIYNRPVSAKRI-GDVIS 541
Db 543 NLLESINNRKD-----NTLPLELKGRIYSEIYN--ISAPNTPGSTLI 584
Qy 542 VS-----NC-----IIVDQTSVSLHKSRLLSASDEKCFRPP 574
Db 585 ISWGRKESGFTLRYRSEEARNOWEKCLDLKTNEMNKQIHKLR-----DSOSSEN 637
Qy 575 VTFKFMNDSTIYKQGLGVNNEILLTTT---YLETQCENTEYFQAQTDVYIKNYEHLKT 631
Db 638 -----TDDSAIY-DYTGISTSPVNQSTQQQYDHRGSHSRHSSSTLSMMKN-NRVKS 690
Qy 632 VPLSSI-----ALNFTLLENVDFKVE-----LYTRDEKRLSN 671
Db 691 GDLRSRISSTTLDSPSNLNGSPNTNPSLMSSDATKTPTPDVAIKLYKSTE--LSE 748
Qy 672 VFDIETMFREYNYAQRV-----SGLRKDLDLSTNRNQFVDAFGSLM-----DDLGA 720
Db 749 PLIVNAQI-EYNDLLQKIISQIITSNLVAD--DVNISRLRYKDDGDFVNLNSDDDWGLV 805

RESULT 39
US-10-054-399A-29
; Sequence 29, Application US/10054399A
; Publication No. US20030175712A1
; GENERAL INFORMATION:
; APPLICANT: Arkowitz, Robert A
; TITLE OF INVENTION: NUCLEOTIDE SEQUENCES AND PROTEIN
; FILE REFERENCE: DYOUL3.1A2CP1
; CURRENT APPLICATION NUMBER: US/10/054.399A
; CURRENT FILING DATE: 2002-12-13
; PRIOR FILING DATE: 1998-10-08
; PRIOR FILING DATE: 1998-10-08
; PRIOR FILING DATE: 1997-10-15
; PRIOR FILING DATE: 1997-10-15
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: US 09/732,180
; PRIOR FILING DATE: 2000-12-07
; PRIOR APPLICATION NUMBER: US 60/169,699
; PRIOR FILING DATE: 1999-12-07
; PRIOR APPLICATION NUMBER: PCT/GB98/03033
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 29
; LENGTH: 813
; TYPE: PRT
; ORGANISM: Candida albicans
US-10-054-399A-29

Query Match 2.5%; Score 115; DB 12; Length 813;
Best Local Similarity 19.2%; Pred. No. 1.2;
Matches 150; Conservative 107; Mismatches 267; Indels 256; Gaps 40;

Qy 75 NIYGSPSTFPYRVCASGVDFRFTDHCVPDASDMVHSEGLLIYKQNIIPMFVRK 134
Db 148 NISVDPAATHLWKLFQO---GAPFCVLINHLIPDSQIPVWSSDDLCKKSVYDFLIAVKT 204
Qy 135 Y-----RKVVTTSTVYNGIYSDSITNQHTFYKSIETPEWETEKMDTIYOCFNSRLNTG 189
Db 205 QLNFDENMFTISNVFSDNAODLI-----KLIQVINKL-----L 238
Qy 190 LTYVDR-----DDINNTVFLQPD-----GVTPDVKRYSGQPELYLEPGWFGYSR 235
Db 239 AEYSDASTSGGGEDVNDV--QITDERSKVFREIETERYKYVDLELMCK-----Y 288

Qy 236 RRTTVNCELMD-----MEARSNPPPDF-----FVTATGDT 265
Db 289 RQDLIAEALNSSQIHLFPNLNEIIDFORFLNGLECNINPIRYQIRGSVFIHAS--- 345
Qy 266 VEMSPFWSGDEDDHENKHEKPFVSVINNYKVVDYQNRGTVPGLKTRIFLD-----REBYT 321
Db 346 --LGPENAYE-----PW---TIGQLTAIDLINKEAANKKSSSLDDPGFELQSYI 390
Qy 322 LSWEKHLKMSYCPPLNLWAFYN-----GIQTEHSGSYHFVANDITASFTTSKEDMKEFNT 377
Db 391 L---KPIQRLCKYPLLLKELIKTSPESYKQDPHGSSSTSFNELLVAKTAMKELANQVN- 446
Qy 378 TYHCLNEELKAELEKKYKAVNTHSKYGDLYKFTDGLLYVWQPLIQLNLLDAKQNLNN 437
Db 447 -----BAORRAENIEHLEK---ERVGNWRGNLDAQOGLLPHG--QVGVKAENEKEY 496
Qy 438 ETY-----SRRSROAESTTDPMMEMTNGAGGEYSSSENSITVAQVQYAYD 483
Db 497 VAYLFEKIVFFFTTIDTKSDQKEKKS-----KFSTRKRSTSSNLSSSTT 542
Qy 484 NLRIRINILEDLSKAWCREQHRAALVWNELS-KINPTSVMSMIYNRPVSAKRI-GDVIS 541
Db 543 NLLESINNRKD-----NTLPLELKGRIYSEIYN--ISAPNTPGSTLI 584
Qy 542 VS-----NC-----IIVDQTSVSLHKSRLLSASDEKCFRPP 574
Db 585 ISWGRKESGFTLRYRSEEARNOWEKCLDLKTNEMNKQIHKLR-----DSOSSEN 637
Qy 575 VTFKFMNDSTIYKQGLGVNNEILLTTT---YLETQCENTEYFQAQTDVYIKNYEHLKT 631
Db 638 -----TDDSAIY-DYTGISTSPVNQSTQQQYDHRGSHSRHSSSTLSMMKN-NRVKS 690
Qy 632 VPLSSI-----ALNFTLLENVDFKVE-----LYTRDEKRLSN 671
Db 691 GDLRSRISSTTLDSPSNLNGSPNTNPSLMSSDATKTPTPDVAIKLYKSTE--LSE 748
Qy 672 VFDIETMFREYNYAQRV-----SGLRKDLDLSTNRNQFVDAFGSLM-----DDLGA 720
Db 749 PLIVNAQI-EYNDLLQKIISQIITSNLVAD--DVNISRLRYKDDGDFVNLNSDDDWGLV 805

RESULT 40
US-09-732-180-2
; Sequence 2, Application US/09732180
; Patent No. US20020137702A1
; GENERAL INFORMATION:
; APPLICANT: Arkowitz, Robert A
; APPLICANT: Nern, Peter MA
; TITLE OF INVENTION: Yeast receptor
; FILE REFERENCE: DYO25.001AUS
; CURRENT APPLICATION NUMBER: US/09/732,180
; CURRENT FILING DATE: 2000-12-07
; PRIOR APPLICATION NUMBER: US 60/169,699
; PRIOR FILING DATE: 1999-12-07
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 844
; TYPE: PRT
; ORGANISM: Candida albicans
US-09-732-180-2

Query Match 2.5%; Score 115; DB 10; Length 844;
Best Local Similarity 19.2%; Pred. No. 1.3;
Matches 150; Conservative 107; Mismatches 267; Indels 256; Gaps 40;

Qy 75 NIYGSPSTFPYRVCASGVDFRFTDHCVPDASDMVHSEGLLIYKQNIIPMFVRK 134
Db 162 NISVDPAATHLWKLFQO---GAPFCVLINHLIPDSQIPVWSSDDLCKKSVYDFLIAVKT 218
Qy 135 Y-----RKVVTTSTVYNGIYSDSITNQHTFYKSIETPEWETEKMDTIYOCFNSRLNTG 189

```

Db 219 QLNFDENFTISNFSDNAQDLI-----KIIVINKL-----L 252
QY 190 LTYVDR-----DDINMTVFLQVD-----GVTPDVKRYGSOPELYLEPGWFGSYR 235
Db 253 AEYSDASDSGGDEDVMDV--QITDERSKVFREIETERYKVQDLELMCK-----Y 302
QY 236 RRTTVNCELMD-----MFARSNPPDF-----FVTATGDT 265
Db 303 RODLIEAENLSSEQIHLLFPNLNEIIDFQRRFLNGLECNINVPRIYQIGSVFIHAS--- 359
QY 266 VEMSPWSEDDHENKHEKDFVSVNNYKVVQNRGTVPGLKTRIFLD-----REYVT 321
Db 360 --LGPFNAYE-----PW---TIGQLTAIDLINKEAANKKSSSLDPGFELOSQYI 404
QY 322 LSWKHLKNMYSYCPILTWKAFYN---GIQTEHSGSYHFVANDITASFTTSKEDMKPEPT 377
Db 405 L---KPIORLCKYPLLKELIKTSPEYSKODPHGSSSTSFNELLVAKTAMKELANQVN- 460
QY 378 TYHCLNEEBIKAEIKKYAKVNSTHSGYKDYKFTDGGLYLVWQPLIQNRLLDANKLNN 437
Db 461 ----EAOORAEINIEHLEKLK---ERVGNWRGFNLDAQGELLFHG--QVGKDAENEKEY 510
QY 438 ETY-----SRRSRROAESTTDPMMEMTNGAGGEYSSENSITVAQVQYAYD 483
Db 511 VAYLFKIVFFTEIDDTKKSDQKEKKS-----KFSTRKRSTSSNLSSTT 556
QY 484 NLRIRINNILEDLSKAWCREQRAALVMNELS-KINPTSVMSMIYNRPVSAKRI-GDVIS 541
Db 557 NLLESINNSRKD-----NTLPLELKGRVYISEIYN--ISAPNTPGSLTI 598
QY 542 VS-----INVDQTSVSLHKSURLLSASDEKCFSRPP 574
Db 599 ISWSGRKESGFTLRYRSEARNQWEKCLRDLTNEMNKQIHKKLK----DSDSFN--- 651
QY 575 VTFKPMNDSTIYKQGVNNEILLTTT---YLETQCENTEYVFOAKTDMYIKNYEHLKT 631
Db 652 ----TDDSAIY-DYTGISTSPVNSTOQQYDHRGSHSRHSSSTLSMMKN-NRVKS 704
QY 632 VPLSSI-----TTLDTFI-----ALNFTLLENVDFKVE-----LYTRDEKRLSN 671
Db 705 GDLRSISSTSTLDSFNNLNGSPNTNPSLMSSTATKTIPTFDVAIKLLYKSTB--LSE 762
QY 672 VFDIETMFREYNYAQRV-----SGLRKDLDDLSTNRNQFVDAFGSLM-----DDLGA 720
Db 763 PLIVNAQI-EYNDLLQKIISQIITSNLVAD--DVNISRLRYKDDGDFVNLNSDDDWGLV 819

```

RESULT 41

```

US-10-054-399A-24
; Sequence 24, Application US/10054399A
; Publication No. US2003017512A1
; GENERAL INFORMATION:
; APPLICANT: Arkowitz, Robert A
; APPLICANT: Kern, Peter MA
; TITLE OF INVENTION: NUCLEOTIDE SEQUENCES AND PROTEIN
; FILE OF INVENTION: SEQUENCES
; FILE REFERENCE: DY0U13.1A2CP1
; CURRENT APPLICATION NUMBER: US/10/054,399A
; CURRENT FILING DATE: 2002-12-13
; PRIOR APPLICATION NUMBER: 09/168,474
; PRIOR FILING DATE: 1998-10-08
; PRIOR APPLICATION NUMBER: 08/951,141
; PRIOR FILING DATE: 1997-10-15
; PRIOR APPLICATION NUMBER: 09/529,106
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: US 09/732,180
; PRIOR FILING DATE: 2000-12-07
; PRIOR APPLICATION NUMBER: US 60/169,699
; PRIOR FILING DATE: 1999-12-07
; PRIOR APPLICATION NUMBER: PCT/GB98/03033
; PRIOR FILING DATE: 1998-10-08
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: FastSeq for Windows Version 4.0

```

```

; SEQ ID NO 24
; LENGTH: 844
; TYPE: PRT
; ORGANISM: Candida albicans
US-10-054-399A-24

```

```

Query Match 2.5%; Score 115; DB 12; Length 844;
Best Local Similarity 19.2%; Pred. No. 1.3;
Matches 150; Conservative 107; Mismatches 267; Indels 256; Gaps 40;

```

```

QY 75 NIYSPSTFPYRVCSAGVDFRQTDHVCPSDASDMVHSEGIILYKONIIIPMFVRVK 134
Db 162 NISVDPAHLKFLQQ---GAPFCVLIINHILDSQIPVSSDDLRIKSKSVYDFLIAVKT 218
QY 135 Y----RKVATTTSTVNGIYSDSTINQHTFYKSIETPETERKMDTIYOCFNSLRNLGTGNL 189
Db 219 QLNFDENFTISNFSDNAQDLI-----KIIVINKL-----L 252
QY 190 LTYVDR-----DDINMTVFLQVD-----GVTPDVKRYGSOPELYLEPGWFGSYR 235
Db 253 AEYSDASDSGGDEDVMDV--QITDERSKVFREIETERYKVQDLELMCK-----Y 302
QY 236 RRTTVNCELMD-----MFARSNPPDF-----FVTATGDT 265
Db 303 RODLIEAENLSSEQIHLLFPNLNEIIDFQRRFLNGLECNINVPRIYQIGSVFIHAS--- 359
QY 266 VEMSPWSEDDHENKHEKDFVSVNNYKVVQNRGTVPGLKTRIFLD-----REYVT 321
Db 360 --LGPFNAYE-----PW---TIGQLTAIDLINKEAANKKSSSLDPGFELOSQYI 404
QY 322 LSWKHLKNMYSYCPILTWKAFYN---GIQTEHSGSYHFVANDITASFTTSKEDMKPEPT 377
Db 405 L---KPIORLCKYPLLKELIKTSPEYSKODPHGSSSTSFNELLVAKTAMKELANQVN- 460
QY 378 TYHCLNEEBIKAEIKKYAKVNSTHSGYKDYKFTDGGLYLVWQPLIQNRLLDANKLNN 437
Db 461 ----EAOORAEINIEHLEKLK---ERVGNWRGFNLDAQGELLFHG--QVGKDAENEKEY 510
QY 438 ETY-----SRRSRROAESTTDPMMEMTNGAGGEYSSENSITVAQVQYAYD 483
Db 511 VAYLFKIVFFTEIDDTKKSDQKEKKS-----KFSTRKRSTSSNLSSTT 556
QY 484 NLRIRINNILEDLSKAWCREQRAALVMNELS-KINPTSVMSMIYNRPVSAKRI-GDVIS 541
Db 557 NLLESINNSRKD-----NTLPLELKGRVYISEIYN--ISAPNTPGSLTI 598
QY 542 VS-----INVDQTSVSLHKSURLLSASDEKCFSRPP 574
Db 599 ISWSGRKESGFTLRYRSEARNQWEKCLRDLTNEMNKQIHKKLK----DSDSFN--- 651
QY 575 VTFKPMNDSTIYKQGVNNEILLTTT---YLETQCENTEYVFOAKTDMYIKNYEHLKT 631
Db 652 ----TDDSAIY-DYTGISTSPVNSTOQQYDHRGSHSRHSSSTLSMMKN-NRVKS 704
QY 632 VPLSSI-----TTLDTFI-----ALNFTLLENVDFKVE-----LYTRDEKRLSN 671
Db 705 GDLRSISSTSTLDSFNNLNGSPNTNPSLMSSTATKTIPTFDVAIKLLYKSTB--LSE 762
QY 672 VFDIETMFREYNYAQRV-----SGLRKDLDDLSTNRNQFVDAFGSLM-----DDLGA 720
Db 763 PLIVNAQI-EYNDLLQKIISQIITSNLVAD--DVNISRLRYKDDGDFVNLNSDDDWGLV 819

```

RESULT 42

```

US-09-815-242-13693
; Sequence 13693, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.

```


Db 1465 ESVAESVASESVASVAVSDISLDTTSEWSTSDSKIVASTVSPSEORSSIPMS 1524
Qy 395 AKVNSTHSGYGLKFKYKTDGGLYLWQPLIQNRLDANKLNNE-----TYSRR----- 443
Db 1535 SDESSESR-----ESSGTL-----SENSDSIPTTFSTRYSPSGM 1563
Qy 444 -SRQAESITDPMEMTNGAGGEYSSENSITVAQVQAYDNLRIINNILEDLSKAWCR 502
Db 1564 SSRHVTNSTETSVSDWSSVAGDTESSVSV-----ISESESPTS 1606
Qy 503 EQHRAALVWNLKSKINPTSVSMIYNRPVSAKRIGDVISVNCIVVDQTSV---SLHKS 559
Db 1607 ESVAESVASE--SVTAVSDISLDT-----TSEVSTSDSKIVPSTVSPSSQRSS 1657
Qy 560 RLLSASDEKCFSPRPVTFKPMNDSTIYKGLGVNNEILLTTTLETCQENTETYYFQAKTD 619
Db 1658 PMSSDSESSRE-----SSGTL-----LSENSDSIPTTF-----STRYSPSGMS 1701
Qy 620 MYIYKN 625
Db 1702 SRHYTN 1707

RESULT 44
US-10-304-095-6
; Sequence 6, Application US/10304095
; Publication No. US20030134275A1
; GENERAL INFORMATION:
; APPLICANT: Long, David M.
; APPLICANT: Metz, Anneke M.
; APPLICANT: Love, Ruschelle A.
; FILE OF INVENTION: Telomerase Reverse Transcriptase (TERT) Genes
; CURRENT APPLICATION NUMBER: US/10/304,095
; CURRENT FILING DATE: 2002-11-26
; PRIOR APPLICATION NUMBER: US/09/417,485
; PRIOR FILING DATE: 2002-06-14
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 6
; LENGTH: 2184
; TYPE: PRT
; ORGANISM: Plasmodium falciparum
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (330)...(335)
; OTHER INFORMATION: Xaa at position 330 = Leu or Ile;
; OTHER INFORMATION: Xaa at position 335 = Asp or Gly.

US-10-304-095-6

Query Match 2.5%; Score 114; DB 12; Length 2184;
Best Local Similarity 17.6%; Pred. No. 7.1; Mismatches 297; Indels 338; Gaps 46;
Matches 166; Conservative 140;
Qy 61 RGNNEORSEQKNYISPSPTFYRVCSASGVDFRQTDHVCPCDASDMVHSEGLLI 120
Db 1031 RNNNNKKAQNEKNI-----DSDNLEKKKKKIYI 1061
Qy 121 YK-QNIIP---PMFRV-----RKVRKVVTSTV-----YNGI 148
Db 1062 YKIKNIIEKRNPMKLNSINHFISKKLRINWIPKKGLRPLNLTSLNVPPIVKQRIPEI 1121
Qy 149 YSDSITNOHTFYKSIPEPETERKMDT-----IYQCFNSRLN 184
Db 1122 LKSKSSSEFYFNILNUNLREKDKNIKKRKYKNKFNPNVSLNINCFSLKCLGNMHN 1181
Qy 185 TGG---NLLTYDVRDINNVTFLQPVGVTPDVKRYGSOPELYLEPGFWGSYRRRTVN 241
Db 1182 NNSLFKNLTGTELELKKWLH-----YLK-NWFKYKKRKKYIK 1222
Qy 242 CELMDMFARSNPFFDFTATGDTVMGPFWSGEDHKNKHEKPW-----FVSVINNYK 296
Db 1223 NKL-----KNNKIYAYI-CIGD-----FSNCYEHLN--HNYLFKILKNFFDNINNF 1267

Qy 297 VVDYQNRGTVPGLKTRIFLDREEVTLSEKHLKN--MSYCPLTLWKAPYNG----- 345
Db 1268 F-----IYLPKRSFRL-YNNKLNNSFLSYYPVVV-KSF--GLHYIRNLRE 1308
Qy 346 --IQTEHSGSVHFVANDITASFTTSKEDMKEFNTYHCLNBEIEKAEIEK----- 392
Db 1309 LIIKSHLNDNHHFLNQMPK--TKSKSDLYFADSYKSL-----QVDRKDFMTIITVI 1360
Qy 393 KYAKVNSTHSGYGLKFKYKTDGGLYLWQPLIQNRLDANKLNNETYSSRRSQAESFT 452
Db 1361 RYYVLNI-----YFSIKEFKLNRKNIFYQ-----IFQENQMGVYLSVRDKKRVENIK 1409
Qy 453 DPMEMTNGAGGEYSSENSITVAQVQAYDNLRIINN-----ILEDLKAWC 501
Db 1410 KWL-----NSMKINHDIELESKNSINNNKFMICTNHEQDTEKGNQTN 1458
Qy 502 REQHRAAL--VMNELSKINPTSVSMIYNRPVSAKRIGDVISVNCIVVDQTSVSLHK-- 557
Db 1459 KEKHDIVIGPIYNNSPDSTTTTHSSNNYK-----GNNIHVSGDYKNDGL--LHKGN 1507
Qy 558 -SLRLLSASDEKCFSPRPVTFKPMNDSTIYKGLGVNNEILLTTTLE--TCQENTETYYFQ 615
Db 1508 NSMNECYVKDICKNN-----NNNNNNNNNNNN-----SYNKLNCVTN-----N 1548
Qy 616 AKTDMYIKNVEHLKTVPLSSITLDTFIALNFTLLENVDFKVELYTRDEKRLSNVFDI 675
Db 1549 SKNDIIKYHK-----TIDTNSKNHTYFKNKLNF-----DKIISNIYGL 1590
Qy 676 ETMPREYN-----YYA-----QVSGLRKOL-----LDLSTNRNQFVDA 709
Db 1591 PQGFSLSNLCSLYAYLDKNEEFQNLSEKQINNKYFLANGTCYFNLSLILRIFDD 1650
Qy 710 FGLMDDLGAQGVTVNAVSGVATLPSISIVTGFNFINKFPGMLMIIVIGULFAIYFL 769
Db 1651 F-----LFTLNKKNIKIFKN-----LLKKTKTWSNIN 1679
Qy 770 TKTKIYETAPI---KWIYPEIDKLKEREKSEIAPISEEELERIVLAMHIHQO---N 821
Db 1680 SSKTKIPRIPLYKNDLLIYFNQYKQKKYK-----IKKKKIQSVNRKRIHQLNVAN 1734
Qy 822 SHMETKTRKDPKDSILTRAQNMRLKRSGY---SNLKNAESV 859
Db 1735 KKHSTSVQKDKINKYINLIHPTIQKNDSDVLSNSIMNFERI 1775

RESULT 45
US-10-242-056-53
; Sequence 53, Application US/10242056
; Publication No. US20030113323A1
; GENERAL INFORMATION:
; APPLICANT: Ensign, Jerald C
; APPLICANT: Bowen, David J
; APPLICANT: Petell, James
; APPLICANT: Fatig, Raymond
; APPLICANT: Schoonover, Sue
; APPLICANT: French-Constant, Richard
; APPLICANT: Orr, Gregory L
; APPLICANT: Merlo, Donald J
; APPLICANT: Roberts, Jean L
; APPLICANT: Rocheleau, Thomas A
; TITLE OF INVENTION: Insecticidal Protein Toxins from
; TITLE OF INVENTION: Photorhabdus
; NUMBER OF SEQUENCES: 88
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dowelanco
; STREET: 9330 Zionsville Road
; CITY: Indianapolis
; STATE: IN
; COUNTRY: US
; ZIP: 46268
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk

```
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/242,056
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/063,615
FILING DATE: 18-MAY-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/395,497
FILING DATE: 28-FEB-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/007,255
FILING DATE: 06-NOV-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/608,423
FILING DATE: 28-FEB-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/705,484
FILING DATE: 28-AUG-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/743,699
FILING DATE: 06-NOV-1996
ATTORNEY/AGENT INFORMATION:
NAME: Borucki, Andrea T.
REGISTRATION NUMBER: 33651
REFERENCE/DOCKET NUMBER: 50301E
TELEPHONE: 317-337-4846
TELEPHONE: 317-337-4847
INFORMATION FOR SEQ ID NO: 53:
SEQUENCE CHARACTERISTICS:
LENGTH: 1844 amino acids
TYPE: amino acids
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-10-242-056-53

Query Match      2.5%; Score 113.5; DB 15; Length 1844;
Best Local Similarity 19.5%; Pred. No. 6;
Matches 119; Conservative 82; Mismatches 219; Indels 189; Gaps 34;

Qy      108 ASDMVHSEGLLIYKQN----IIPFMRVRKYRKVTTSTVNGIY-----SDSITN-Q 156
Db      1118 STDAESGLGYCTGYQGEDTLVNFYSMQSSYSYTDNNAPVTGLYIFADMSSDNMTNAQ 1177
Qy      157 HTFYKSIETPETE--KMDTI-----YOCFNSLRLAT---GG 187
Db      1178 ATNY-----WNNSYPQEDTVADPDSDNKKVITRRVNNRYAEDYEIPSSVTSNYSYSGD 1232
Qy      188 NLITYVDRDDINMTVFLQPDVGTPDVKRYGSOPELYLEPGW---FWGSYRRRTTVNCE 243
Db      1233 HSLTML-----YGGSVNITFESAEDLRLSTNMAISIIHNGYAGTRRIQCN 1279
Qy      244 LMDMFARSNPPFPFVATGDT-----VEMSPFWSGE--DD-----HENKQHE-KPWFS 290
Db      1280 LMKQYASLGDKFIYDSFDDANRNLVPLFKFGKDSNDDSDSICIYENENPSSDDKKWYFS 1339
Qy      291 VINNYKVVDYQNRCT--VPLGKTRIFLDREYTLSEWEHLKNMSYCPPLTLWKAFYNGIQT 348
Db      1340 SKDKNKTADY-NGTQCIDAGTSN-----KDFYNNLQ- 1370
Qy      349 EHGSGYHFVANDIPASFTTSKEDMKFEFTTYHCLNEEIKAEIKYAKVNSTHSGYDGLK 408
Db      1371 -----EIEVISVTGGVWSSYKISNPIN-----INTGIDS-----AKVKVTVKAGDDQ 1413
Qy      409 YEKTDGGLYLWQPL-----IQNRLDAKKK--LNNE-----TYSRRSRROAES 450
Db      1414 IFTADNSTYVPQAPSPSEEMIVQFNLLTIDCKNLNFDINQAHIEIDFTATAQDGRFLGA 1473
```

```
Qy      451 TTDPMMEMTCNGAGGE-----YSSENSITVAQVOYAYDNLIRINNILEDLSKAWCREQH 505
Db      1474 ETP-IIPVTYKVLGTENVIALYSENGVQVMOIG-AY---RPLNLTIFAQ----- 1518
Qy      506 RAALVWNELSKINP--TSVMSMIYNRPVSARKRIGDVISVSNCIIVDPQTSVSLHKSURLLS 563
Db      1519 -----QLVSRANRGIDAVLSM-ETQNIQEPQLGAGTYVQ--LVLDKYDESIH----- 1562
Qy      564 ASDEKCFSRPPVTFKFMNDS-TIYKQOLGVNNEILL---TTTYLETQCENTEYFFQAK-- 617
Db      1563 -GTNKSFAIEYVDIFKENDSFVIYQOGLSETSQTVVVKVFLSYFIEATGNKNHLMVRAKYQ 1621
Qy      618 ---TDMYIY 623
Db      1622 KETTDKILF 1630
```

Search completed: October 28, 2003, 15:41:52
Job time : 94 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 28, 2003, 15:29:40 ; Search time 392 Seconds
(without alignments)
2007.856 Million cell updates/sec

Title: US-10-055-364-24

Perfect score: 4547

Sequence: 1 MAGSLKRGSLVLMYLYQV.....KRGYSNLKNAESVEMLNTL 865

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 5728757 seqs, 909918778 residues

Total number of hits satisfying chosen parameters: 5728757

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Pending Patents AA Main:*

- 1: /cgn2_6/ptodata/2/paa/pctus COMB.pcp.*
- 2: /cgn2_6/ptodata/2/paa/US06 COMB.pcp.*
- 3: /cgn2_6/ptodata/2/paa/US07 COMB.pcp.*
- 4: /cgn2_6/ptodata/2/paa/US08 COMB.pcp.*
- 5: /cgn2_6/ptodata/2/paa/US08 COMB.pcp.*
- 6: /cgn2_6/ptodata/2/paa/US08 COMB.pcp.*
- 7: /cgn2_6/ptodata/2/paa/US08 COMB.pcp.*
- 8: /cgn2_6/ptodata/2/paa/US08 COMB.pcp.*
- 9: /cgn2_6/ptodata/2/paa/US08 COMB.pcp.*
- 10: /cgn2_6/ptodata/2/paa/US08 COMB.pcp.*
- 11: /cgn2_6/ptodata/2/paa/US08 COMB.pcp.*
- 12: /cgn2_6/ptodata/2/paa/US08 COMB.pcp.*
- 13: /cgn2_6/ptodata/2/paa/US08 COMB.pcp.*
- 14: /cgn2_6/ptodata/2/paa/US08 COMB.pcp.*
- 15: /cgn2_6/ptodata/2/paa/US08 COMB.pcp.*
- 16: /cgn2_6/ptodata/2/paa/US08 COMB.pcp.*
- 17: /cgn2_6/ptodata/2/paa/US08 COMB.pcp.*
- 18: /cgn2_6/ptodata/2/paa/US08 COMB.pcp.*
- 19: /cgn2_6/ptodata/2/paa/US08 COMB.pcp.*
- 20: /cgn2_6/ptodata/2/paa/US08 COMB.pcp.*
- 21: /cgn2_6/ptodata/2/paa/US08 COMB.pcp.*
- 22: /cgn2_6/ptodata/2/paa/US08 COMB.pcp.*
- 23: /cgn2_6/ptodata/2/paa/US08 COMB.pcp.*
- 24: /cgn2_6/ptodata/2/paa/US08 COMB.pcp.*
- 25: /cgn2_6/ptodata/2/paa/US08 COMB.pcp.*
- 26: /cgn2_6/ptodata/2/paa/US08 COMB.pcp.*
- 27: /cgn2_6/ptodata/2/paa/US08 COMB.pcp.*
- 28: /cgn2_6/ptodata/2/paa/US08 COMB.pcp.*
- 29: /cgn2_6/ptodata/2/paa/US08 COMB.pcp.*
- 30: /cgn2_6/ptodata/2/paa/US08 COMB.pcp.*
- 31: /cgn2_6/ptodata/2/paa/US08 COMB.pcp.*
- 32: /cgn2_6/ptodata/2/paa/US08 COMB.pcp.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match Length	ID	Description
1	4547	100.0	865	1 PCT-US00-18647-24

2	4547	100.0	865	26	US-10-055-364-24	Sequence 24, Appl
3	2473	54.4	793	26	US-10-055-364-55	Sequence 55, Appl
4	2438	53.6	830	26	US-10-055-364-45	Sequence 45, Appl
5	1953.5	43.0	874	12	US-08-804-439-16	Sequence 16, Appl
6	1953.5	43.0	874	17	US-09-301-390-16	Sequence 16, Appl
7	1953.5	43.0	874	17	US-09-338-326-16	Sequence 16, Appl
8	1949.5	42.9	808	12	US-08-804-439-14	Sequence 14, Appl
9	1949.5	42.9	808	17	US-09-301-390-14	Sequence 14, Appl
10	1949.5	42.9	808	17	US-09-338-326-14	Sequence 14, Appl
11	1940	42.7	824	26	US-10-055-364-43	Sequence 43, Appl
12	1937	42.6	845	12	US-08-804-439-94	Sequence 94, Appl
13	1937	42.6	845	17	US-09-301-390-94	Sequence 94, Appl
14	1937	42.6	845	17	US-09-338-326-94	Sequence 94, Appl
15	1927.5	42.4	829	14	US-09-053-601-3	Sequence 3, Appl
16	1927.5	42.4	829	23	US-09-831-000-13	Sequence 13, Appl
17	1918	42.2	823	26	US-10-055-364-37	Sequence 37, Appl
18	1916	42.1	792	26	US-10-055-364-42	Sequence 42, Appl
19	1902	41.8	808	26	US-10-055-364-38	Sequence 38, Appl
20	1898.5	41.8	791	26	US-10-055-364-41	Sequence 41, Appl
21	1871.5	41.2	818	26	US-10-055-364-44	Sequence 44, Appl
22	1848.5	40.7	857	7	US-03-360-107-113	Sequence 113, Appl
23	1848.5	40.7	857	12	US-08-804-439-18	Sequence 18, Appl
24	1848.5	40.7	857	17	US-09-301-390-18	Sequence 18, Appl
25	1848.5	40.7	857	17	US-09-338-326-18	Sequence 18, Appl
26	1844.5	40.6	829	26	US-10-055-364-46	Sequence 46, Appl
27	1839	40.4	856	8	US-08-471-913-103	Sequence 103, Appl
28	1839	40.4	856	8	US-08-475-668-103	Sequence 103, Appl
29	1839	40.4	856	8	US-08-484-223-103	Sequence 103, Appl
30	1839	40.4	856	8	US-08-484-223A-103	Sequence 103, Appl
31	1839	40.4	856	8	US-08-484-223D-103	Sequence 103, Appl
32	1839	40.4	856	8	US-08-485-551-103	Sequence 103, Appl
33	1839	40.4	856	8	US-08-487-266-103	Sequence 103, Appl
34	1839	40.4	856	8	US-08-487-266A-103	Sequence 103, Appl
35	1839	40.4	856	8	US-08-487-355-103	Sequence 103, Appl
36	1839	40.4	856	8	US-08-487-355A-103	Sequence 103, Appl
37	1839	40.4	856	13	US-08-519-600-103	Sequence 103, Appl
38	1839	40.4	856	19	US-09-502-445-103	Sequence 103, Appl
39	1839	40.4	856	28	US-10-267-682-103	Sequence 103, Appl
40	1839	40.4	856	28	US-10-267-748-103	Sequence 103, Appl
41	1792	39.4	849	12	US-08-804-439-17	Sequence 17, Appl
42	1792	39.4	849	17	US-09-301-390-17	Sequence 17, Appl
43	1792	39.4	849	17	US-09-338-326-17	Sequence 17, Appl
44	1787	39.3	874	12	US-08-804-439-15	Sequence 15, Appl
45	1787	39.3	874	17	US-09-301-390-15	Sequence 15, Appl

ALIGNMENTS

RESULT 1
PCT-US00-18647-24
; Sequence 24, Application PC/TUS0018647
; GENERAL INFORMATION:
; APPLICANT: Patience, Clive
; TITLE OF INVENTION: Swine Gamma Herpesvirus DNA and Methods of Use
; FILE REFERENCE: 61750-299
; CURRENT APPLICATION NUMBER: PCT/US00/18647
; CURRENT FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: U.S. 60/142,736
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: U.S. 60/168,532
; PRIOR FILING DATE: 1999-12-02
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 24
; LENGTH: 865
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:Deduced amino
; OTHER INFORMATION: acid sequence of porcine gamma herpesvirus gpb
; OTHER INFORMATION: gene
PCT-US00-18647-24

Query Match 100.0%; Score 4547; DB 1; Length 865;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 865; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MAGSLKRGSLALWLYQVALYSLAETGVTSPNTATWSTESPLTGHYTHDSSHGE 60
Db 1 MAGSLKRGSLALWLYQVALYSLAETGVTSPNTATWSTESPLTGHYTHDSSHGE 60

Qy 61 RGNENRDSBEONKNIYGSPTFPYRVCASGVGDVFRFQTDHVCDDASDMVHSEGIILLI 120
Db 61 RGNENRDSBEONKNIYGSPTFPYRVCASGVGDVFRFQTDHVCDDASDMVHSEGIILLI 120

Qy 121 YKQNIIPFMRVRYKRVKVTSTVYNGIYSDSIINQHTFYKSIPEWETEKMDTIYOCFNS 180
Db 121 YKQNIIPFMRVRYKRVKVTSTVYNGIYSDSIINQHTFYKSIPEWETEKMDTIYOCFNS 180

Qy 181 LRLNTGNNLLTYVDRDDINMTVFLQPDVGVTPDKRYGSOPELYLEPGFWGYSYRRRTTV 240
Db 181 LRLNTGNNLLTYVDRDDINMTVFLQPDVGVTPDKRYGSOPELYLEPGFWGYSYRRRTTV 240

Qy 241 NCELMDMFARSNPFPDFVTATGDTVEMSPFWSGEDDHENKMKHPWFSVINNYKVVDY 300
Db 241 NCELMDMFARSNPFPDFVTATGDTVEMSPFWSGEDDHENKMKHPWFSVINNYKVVDY 300

Qy 301 QNRGTVPGLKTRIFLDREEVYTLSEKHLKMSYCPLTLWKAFYNGIOTHSGSYHFVAND 360
Db 301 QNRGTVPGLKTRIFLDREEVYTLSEKHLKMSYCPLTLWKAFYNGIOTHSGSYHFVAND 360

Qy 361 ITASFTTSKEDMKFNTTYHCLNEEIKAIEKKYAKVNSTHSGYDLYFKTDGGLYLW 420
Db 361 ITASFTTSKEDMKFNTTYHCLNEEIKAIEKKYAKVNSTHSGYDLYFKTDGGLYLW 420

Qy 421 OPLIQNRLLDKAKNKLNETYSRRSRQAESTTDPMMEMTNGAGGEYSSENSITVAQVQY 480
Db 421 OPLIQNRLLDKAKNKLNETYSRRSRQAESTTDPMMEMTNGAGGEYSSENSITVAQVQY 480

Qy 481 AYDNLRIIRINILEDLSKACREOHRALVWNELSKINPTSVMSMIYNNRPVSAKRIGDVI 540
Db 481 AYDNLRIIRINILEDLSKACREOHRALVWNELSKINPTSVMSMIYNNRPVSAKRIGDVI 540

Qy 541 SVSNCIVVDQTSVSLHSLRLLSASDEKCFSRPPVTFFKFNNDSTIYKQGLGVNNEILLTT 600
Db 541 SVSNCIVVDQTSVSLHSLRLLSASDEKCFSRPPVTFFKFNNDSTIYKQGLGVNNEILLTT 600

Qy 601 TYLETQCENTEYFQAOKTDMYIKNYEHLKTVPLSSITTLDTFIALNFTLLENVDFKVIE 660
Db 601 TYLETQCENTEYFQAOKTDMYIKNYEHLKTVPLSSITTLDTFIALNFTLLENVDFKVIE 660

Qy 661 LYTRDEKRLSNVDFDIETMFREYNYAQRVSGRLKDLDDLSTNRNQFVDAFGSLMDDLGA 720
Db 661 LYTRDEKRLSNVDFDIETMFREYNYAQRVSGRLKDLDDLSTNRNQFVDAFGSLMDDLGA 720

Qy 721 GQTVVNAVSGVATLFSISITGFINFIKNPFGGMLMIIVIGLFAIYFLTKTKIYETAP 780
Db 721 GQTVVNAVSGVATLFSISITGFINFIKNPFGGMLMIIVIGLFAIYFLTKTKIYETAP 780

Qy 781 IKMIYPEIDKLKEREKSEAPISEEBELERIVLAMHIHQNSHMETKRPDKDSILTRA 840
Db 781 IKMIYPEIDKLKEREKSEAPISEEBELERIVLAMHIHQNSHMETKRPDKDSILTRA 840

Qy 841 QNMLRKESGYSLNKAESVEMLNTL 865
Db 841 QNMLRKESGYSLNKAESVEMLNTL 865

RESULT 2

US-10-055-364-24
; Sequence 24, Application US/10055364
; GENERAL INFORMATION:
; APPLICANT: Patience, Clive
; TITLE OF INVENTION: Gamma Herpes Virus DNA and Methods of Use
; FILE REFERENCE: *1750-379

CURRENT APPLICATION NUMBER: US/10/055,364
CURRENT FILING DATE: 2002-01-23
PRIOR APPLICATION NUMBER: US/09/612,204
PRIOR FILING DATE: 2000-07-07
PRIOR APPLICATION NUMBER: US/60/142,736
PRIOR FILING DATE: 1999-07-08
PRIOR APPLICATION NUMBER: US/60/168,532
PRIOR FILING DATE: 1999-12-02
NUMBER OF SEQ ID NOS: 55
SOFTWARE: PatentIn version 3.0
SEQ ID NO 24
LENGTH: 865
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Deduced amino acid sequence of porcine gamma herpesvirus gpB gene
US-10-055-364-24

Query Match 100.0%; Score 4547; DB 26; Length 865;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 865; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MAGSLKRGSLALWLYQVALYSLAETGVTSPNTATWSTESPLTGHYTHDSSHGE 60
Db 1 MAGSLKRGSLALWLYQVALYSLAETGVTSPNTATWSTESPLTGHYTHDSSHGE 60

Qy 61 RGNENRDSBEONKNIYGSPTFPYRVCASGVGDVFRFQTDHVCDDASDMVHSEGIILLI 120
Db 61 RGNENRDSBEONKNIYGSPTFPYRVCASGVGDVFRFQTDHVCDDASDMVHSEGIILLI 120

Qy 121 YKQNIIPFMRVRYKRVKVTSTVYNGIYSDSIINQHTFYKSIPEWETEKMDTIYOCFNS 180
Db 121 YKQNIIPFMRVRYKRVKVTSTVYNGIYSDSIINQHTFYKSIPEWETEKMDTIYOCFNS 180

Qy 181 LRLNTGNNLLTYVDRDDINMTVFLQPDVGVTPDKRYGSOPELYLEPGFWGYSYRRRTTV 240
Db 181 LRLNTGNNLLTYVDRDDINMTVFLQPDVGVTPDKRYGSOPELYLEPGFWGYSYRRRTTV 240

Qy 241 NCELMDMFARSNPFPDFVTATGDTVEMSPFWSGEDDHENKMKHPWFSVINNYKVVDY 300
Db 241 NCELMDMFARSNPFPDFVTATGDTVEMSPFWSGEDDHENKMKHPWFSVINNYKVVDY 300

Qy 301 QNRGTVPGLKTRIFLDREEVYTLSEKHLKMSYCPLTLWKAFYNGIOTHSGSYHFVAND 360
Db 301 QNRGTVPGLKTRIFLDREEVYTLSEKHLKMSYCPLTLWKAFYNGIOTHSGSYHFVAND 360

Qy 361 ITASFTTSKEDMKFNTTYHCLNEEIKAIEKKYAKVNSTHSGYDLYFKTDGGLYLW 420
Db 361 ITASFTTSKEDMKFNTTYHCLNEEIKAIEKKYAKVNSTHSGYDLYFKTDGGLYLW 420

Qy 421 OPLIQNRLLDKAKNKLNETYSRRSRQAESTTDPMMEMTNGAGGEYSSENSITVAQVQY 480
Db 421 OPLIQNRLLDKAKNKLNETYSRRSRQAESTTDPMMEMTNGAGGEYSSENSITVAQVQY 480

Qy 481 AYDNLRIIRINILEDLSKACREOHRALVWNELSKINPTSVMSMIYNNRPVSAKRIGDVI 540
Db 481 AYDNLRIIRINILEDLSKACREOHRALVWNELSKINPTSVMSMIYNNRPVSAKRIGDVI 540

Qy 541 SVSNCIVVDQTSVSLHSLRLLSASDEKCFSRPPVTFFKFNNDSTIYKQGLGVNNEILLTT 600
Db 541 SVSNCIVVDQTSVSLHSLRLLSASDEKCFSRPPVTFFKFNNDSTIYKQGLGVNNEILLTT 600

Qy 601 TYLETQCENTEYFQAOKTDMYIKNYEHLKTVPLSSITTLDTFIALNFTLLENVDFKVIE 660
Db 601 TYLETQCENTEYFQAOKTDMYIKNYEHLKTVPLSSITTLDTFIALNFTLLENVDFKVIE 660

Qy 661 LYTRDEKRLSNVDFDIETMFREYNYAQRVSGRLKDLDDLSTNRNQFVDAFGSLMDDLGA 720
Db 661 LYTRDEKRLSNVDFDIETMFREYNYAQRVSGRLKDLDDLSTNRNQFVDAFGSLMDDLGA 720

Qy 721 GQTVVNAVSGVATLFSISITGFINFIKNPFGGMLMIIVIGLFAIYFLTKTKIYETAP 780
Db 721 GQTVVNAVSGVATLFSISITGFINFIKNPFGGMLMIIVIGLFAIYFLTKTKIYETAP 780

Qy	412	TGGGLYVWQPLIQNRLLDAKNKUN-----NETYSRRSRRAQBESTTDPMMWTGNG	462
Db	393	TGGGLLVWQPLVOOQLLDAKGLLDVAKKKQQNTTTTTTRSSRQRSSVS-----SG	444
Qy	463	AGGEYSSENSITVAOVQAYDNLRIRINILEDISKAWCREQHRAALVNNELSKINPTSV	522
Db	445	IDDVYTRASTILLTOQFAYDTLRAQINNVLBELSRAWCREQHRASLMNDELSKINPTSV	504
Qy	523	MSMIYNRPVSAKRGDVISVSGNCIVVDQTSVSLSHKSJRLLSA-SDEKCFSPRPVTTFKFN	581
Db	505	MSSIYGRPVSAKRGDVISVSHCVVDQDSVSLHRSRMVRPGCRDKTHECYSPRPVTTFKFN	564
Qy	582	DSITYKGOLGVANNEILLTTTLYLETCOENTEXYFOAKTDWILYKNYEHLKTVPULSSITTILD	641
Db	565	DSHLHYKGOLGVANNEILLTTTAVETICHENTEHYFOGGNNMYFYKNYRHVKTMPVGDAVUILD	624
Qy	642	TFIALNFILLENVDFPKVIETYELTROBEKLSNVFDIETMFREYNYYAQRVSGLRKDLDLST	701
Db	625	TPMVLNLTLVENIDFQVIELYSREEKRMSTAFDIETMFREYNYYTORVTCGLREDLTDLAT	684
Qy	702	NNRFQVDAPGSLMDDLGAVGTQVNVNAVSGVATLFESSIVTGFIFNFKPNFGGMMLIIWIG	761
Db	685	NRNQFVDAPGSLMDDLGVVGTVLNAVSSVATLFESSIVSGIINFINKPNFGGMMLIFGLIAA	744
Qy	762	VLFPAIVFLTKTKTVETAPIKMVIPEDIKLEREGKSEIAPISEEELERIVLAWHIHQON	821
Db	-745	VVITVILLNRKAKRAFAQNPVQMIPDDIKTTITSOREELOVPISKEHLEDRIMLANMHYHAS	804
Qy	822	SHMETK-----TRDKPKD	834
Db	805	KOPESKODEEOGSITTSGPAD	824

RESULT 5

```

US-08-804-439-16
; Sequence 16, Application US/08804439
; GENERAL INFORMATION:
; APPLICANT: Rose, Timothy M.
; APPLICANT: Rose, Marnix L.
; APPLICANT: Bosch, Kurt
; APPLICANT: Strand, Kurt
; TITLE OF INVENTION: GLYCOPROTEIN B OF THE RFHV/KSHV
; TITLE OF INVENTION: SUBFAMILY OF HERPES VIRUSES
; TITLE OF INVENTION: SUBFAMILY OF HERPES VIRUSES
; NUMBER OF SEQUENCES: 113
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morrison & Foerster
; STREET: 755 Page Mill Road
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304-1018
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION NUMBER: US/08/804,439
; APPLICATION NUMBER: US/08/804,439
; FILING DATE:
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Schiff, J. Michael
; REGISTRATION NUMBER: 40,253
; REFERENCE/DOCKET NUMBER: 29938-20002.20
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 813-5600
; TELEFAX: (415) 494-0792
; TELEX: 706141
; INFORMATION FOR SEQ ID NO: 16:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 874 amino acids
; TYPE: amino acid
; STRANDEDNESS: single

```

```

;
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-804-439-16

Query Match 43.0%; Score 1953.5; DB 12; Length 874;
Best Local Similarity 46.1%; Pred No. 2.2e-174;
Matches 403; Conservative 147; Mismatches 250; Indels 75; Gaps 18

Qy 11 VLALWYLQVALYSLSIAETGVTPPTATWSTSPITLGHYTHDSSHGGRGNENRDS 70
Db 10 VLCLMCV- AALLCGVAQEVVA-----ETTP-----FATH-----RPEVVAE 46

Qy 71 EQNKNIYGSPTFPYRVCSASGV-GDVPFOTDHCPCDASDMVHSEGIILLYKONIPFM 129
Db 10 VLCLMCV- AALLCGVAQEVVA-----ETTP-----FATH-----RPEVVAE 46

Qy 47 ENPAMPF-----LPPRVCGASPTGGEFRFPLEESCPTEDKHIEGIALIYKTIIVPVY 101
Db 10 VLCLMCV- AALLCGVAQEVVA-----ETTP-----FATH-----RPEVVAE 46

Qy 130 FRVKYRKVVTTSTVYNGIYSDISTNOHTFKYSLEPMEKMDTIYOCFNSLRNLNTGNNL 189
Db 10 VLCLMCV- AALLCGVAQEVVA-----ETTP-----FATH-----RPEVVAE 46

Qy 102 FNVKRYRKIMTSTTIYKWSEDAITNOHTRSYAVPLYEVMQMDHYOCFSAQVNEGGHV 161
Db 10 VLCLMCV- AALLCGVAQEVVA-----ETTP-----FATH-----RPEVVAE 46

Qy 190 LTYVDRDDINMTVFLQPDVGTPVDVKRYGSOPELYLEPGFWGYSYRRRTTNCCLMDMFA 249
Db 10 VLCLMCV- AALLCGVAQEVVA-----ETTP-----FATH-----RPEVVAE 46

Qy 162 NTYVDRDGNETAFLKPADGLTSSITRYOSQPEVYATPNLLWSYTRTTTNCVCEVTEMSA 221
Db 10 VLCLMCV- AALLCGVAQEVVA-----ETTP-----FATH-----RPEVVAE 46

Qy 250 RSNPFDPFVATGDTGVMSYFWSGEDDHENKMEKWFVSVINNYKVVDYQNRGTVPGLG 309
Db 10 VLCLMCV- AALLCGVAQEVVA-----ETTP-----FATH-----RPEVVAE 46

Qy 222 RSMKPEFVFTVSGVDITEMSPFLKENGTEPEKILKRPKSLQLLKNYAVTKY----GVGLG 277
Db 10 VLCLMCV- AALLCGVAQEVVA-----ETTP-----FATH-----RPEVVAE 46

Qy 310 K-----TRIFLDREETYTLSEWKLKMSYCPITLWKAFVNGIQTEHSGSYHFVANDITASF 365
Db 10 VLCLMCV- AALLCGVAQEVVA-----ETTP-----FATH-----RPEVVAE 46

Qy 278 QADNATRFPAIFGDISLSMKATTENSSYCDLLLWKGFSAALQTOHNSLHFIANDITASF 337
Db 10 VLCLMCV- AALLCGVAQEVVA-----ETTP-----FATH-----RPEVVAE 46

Qy 366 TTSKEDMKENTYHCLNEETKAEIEKKYAKNTHSKYGDLYKFKTDGGLYLVWQPLIQ 425
Db 10 VLCLMCV- AALLCGVAQEVVA-----ETTP-----FATH-----RPEVVAE 46

Qy 338 STPLSEEAENFTFKCIWNNTQEEITQKLEVEKTRHPNGTKAVKYKTTGNLYIYWQPLIQ 397
Db 10 VLCLMCV- AALLCGVAQEVVA-----ETTP-----FATH-----RPEVVAE 46

Qy 426 NLLDAKKNLNN-----ETYSRRRRQAESTTDPMMEMTNGCAGGYSS--ENS 472
Db 10 VLCLMCV- AALLCGVAQEVVA-----ETTP-----FATH-----RPEVVAE 46

Qy 398 IDLDTAKLYNLNTATASPTSTPTSPRRRRRTSS-----VSGGNGNGNSTKEES 450
Db 10 VLCLMCV- AALLCGVAQEVVA-----ETTP-----FATH-----RPEVVAE 46

Qy 473 ITVAQYQAYDNLRIINILLEDLSKAWCREQHRALVWNLKSKINPTSYMSIYNRPVS 532
Db 10 VLCLMCV- AALLCGVAQEVVA-----ETTP-----FATH-----RPEVVAE 46

Qy 451 VAASQVQFAYDNLRSINRVLGELSAWCAREQYRASLMWYELSKINPTSVMSAIYGRPVS 510
Db 10 VLCLMCV- AALLCGVAQEVVA-----ETTP-----FATH-----RPEVVAE 46

Qy 533 AKRIGDVTSVSNCIYVDQTSVLSKSLRLLSASDEKCFSPRPVTFKPMNDSTIYKCOLGV 592
Db 10 VLCLMCV- AALLCGVAQEVVA-----ETTP-----FATH-----RPEVVAE 46

Qy 511 AKLIGDVSVSDCIISVDQKSFVHKNMK-VPGKEDLCYTRPVVGFKFINGSEIFAGOLGP 569
Db 10 VLCLMCV- AALLCGVAQEVVA-----ETTP-----FATH-----RPEVVAE 46

Qy 593 NNEILLTTLTTCENTYYFOAKTDWYIYKNYEHLKTVPLSGISITTLDTFIALNFTLLE 652
Db 10 VLCLMCV- AALLCGVAQEVVA-----ETTP-----FATH-----RPEVVAE 46

Qy 570 RNEIVLSTQSEVQCHSEHFQAGNQMYKYDYIYVSTNLNLTDIPLTLHTMILTNLSLVE 629
Db 10 VLCLMCV- AALLCGVAQEVVA-----ETTP-----FATH-----RPEVVAE 46

Qy 653 NVDFKVIELYTRDEKRLSNVFDIETMFREYNVYIAQRVSGRLKOLLDD--LSTNRNQVDAFG 711
Db 10 VLCLMCV- AALLCGVAQEVVA-----ETTP-----FATH-----RPEVVAE 46

Qy 630 NIDFKVIELYSKTEKRLSNVFDIETMFREYNVYIAQRVSGRLKOLLDD--LSTNRNQVDAFG 699
Db 10 VLCLMCV- AALLCGVAQEVVA-----ETTP-----FATH-----RPEVVAE 46

Qy 712 SLMDDLGAVQTVNVAHSVATLPSISVTGFIKIPKPFGLMLIIVVI--GVLFPIFYFLT 770
Db 10 VLCLMCV- AALLCGVAQEVVA-----ETTP-----FATH-----RPEVVAE 46

Qy 690 DIMQDLGTIGKVVNVAHSVATLPSISVTGFIKIPKPFGLMLIIVVI--GVLFPIFYFLT 749
Db 10 VLCLMCV- AALLCGVAQEVVA-----ETTP-----FATH-----RPEVVAE 46

Qy 771 KKTIVETAPIKMIYPEIDKLKEREGKSEIAPISEEBELRIVLAWHIHQONSHMETKTRK 830
Db 10 VLCLMCV- AALLCGVAQEVVA-----ETTP-----FATH-----RPEVVAE 46

Qy 750 RRSYIY-SAPIEMLYPGVERAAQEPGAH---PVSEQDIRNILLMGHOFQORQRAEEARR 805
Db 10 VLCLMCV- AALLCGVAQEVVA-----ETTP-----FATH-----RPEVVAE 46

Qy 831 DPK-----DSILTRAQNMLRKR---SGYSNLK 854
Db 10 VLCLMCV- AALLCGVAQEVVA-----ETTP-----FATH-----RPEVVAE 46

Qy 806 EBEVKGKRTLPEVIRDSATSVLRRRGGGQYRLQ 840
Db 10 VLCLMCV- AALLCGVAQEVVA-----ETTP-----FATH-----RPEVVAE 46

```

RESULT 6

APPLICANT: Rose, Timothy M.
APPLICANT: Bosch, Marnix L.
APPLICANT: Strand, Kurt
TITLE OF INVENTION: GLYCOPROTEIN B OF THE RFHV/KSHV
SUBFAMILY OF HERPES VIRUSES
NUMBER OF SEQUENCES: 100
CORRESPONDENCE ADDRESS:
ADDRESSEE: Morrison & Foerster
STREET: 755 Page Mill Road
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304-1018
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/301,390
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/720,229
FILING DATE: 26-SEP-1996
ATTORNEY/AGENT INFORMATION:
NAME: Schiff, J. Michael
REGISTRATION NUMBER: 40,253
REFERENCE/DOCKET NUMBER: 29938-20002.00
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 813-5600
TELEFAX: (415) 494-0792
TELEX: 706141
INFORMATION FOR SEQ ID NO: 16:
SEQUENCE CHARACTERISTICS:
LENGTH: 874 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-301-390-16

Query Match 43.0%; Score 1953.5; DB 17; Length 874;
Best Local Similarity 46.1%; Pred. No. 2,2e-174;
Matches 403; Conservative 147; Mismatches 250; Indels 75; Gaps 18;
QY 11 VLALWYLQVALYSLSIAETGVTSPNTATWSTESPLTGHGYTHDSHGGERGNENRDS 70
DB 10 VLCLMCV--AALLCOGVAQVVA-----ETTP-FAH-----RPEVVAE 46
QY 71 EQKNYGSPTFPYRVCSAGV-GDVERFQTDHVCPSDASDMVHSEGLILYKQNIIPFM 129
DB 47 ENANPF-----LPRVCGASTGGEIPRFLPEECPNTEKDHIEGIALYKTNIVPV 101
QY 130 FVRVYKRVKVTSTVYNGYSDISINQHTFYKSIETPTEKMDTYQCFNSLRNLNLTGNL 189
DB 102 FNVRYKIMSTTIYKGSSEDAINQHTRSVAVPLYEQVMDHYQCFSAVQVNEGHRV 161
QY 190 LTYVDRDINMTVLOPVDGVTDPVKRYGSOPELYLEPFWGFSYRRRTTNCCLMDMFA 249
DB 162 NTTYDRDGNWETAFLKPADGLTSSITRYOSQPEVATPRNLWSYTRRTTNCVETMSA 221
QY 250 RSNPDEFTVATGTVENSFPWSGDEDDHKNHKEKPHFVSVINNYKVVDYQNRGTVP 309
DB 222 RSMKPEFEFTVSGVTIEMSPFKENGTEPEKILKRPHSIQLLKNYAVTKY-----GVGLG 277
QY 310 K-----TRFLDREBYTLWEKHLKMSYCPLTLWKAIFYNGIQTEHSGSYHVFANDITAS 365
DB 278 QADNATRFATGDSLSWKATTENSSYCDLLWKGFSNAQTQNSSLHFIANDITAS 337
QY 366 TTSKEDMKEFNITYHCLNEEIKAEIKRYAKVNSTHSGYGLKYPKTDGGLYVWQPLIQ 425
DB 338 STPLEEAFNFTFKCIWNNTQEEIQKLEKVEKTHRENGTAKVYKTTGNLYVWQPLIQ 397

QY 426 NRLLDKAKNKN-----ETYSRRSRROAESTTDDPMEMTNGAGGEYSS-ENS 472
DB 398 IDLLDTHAKLYNLNATASPTSTPTTSPRRRRRTSS-----VSGGNGNDNSTKEES 450
QY 473 ITVAQVQVAYDNLRIINNILEDLSKAWCREQHRALVNLKSKINPTSVMSIMYNRVPS 532
DB 451 VAASQVQFAYDNLRSINRVLGELSRACRQYRASLMWYELSKINPTSVMSAIYGRVPS 510
QY 533 AKRIGDVISVSNICVVDOTSVLSHLSLRLLSASDEKCFSPPTFKMNDSTIYKQGLGV 592
DB 511 AKLIGDVVSVSDCISVDOKSVFVHKNMK-VPGKEDLCYTRFVVGFKFNGSELFAGQLGP 569
QY 593 NNEILLATYLETQENTYVFOAKTDWYIKNYEHLKTVPLSSITLDTFIALNFTLLE 652
DB 570 RNEIVLSTSQVEVCHSEHYFOAGNQYKYDYVYSTNLNLTPTLHTMITLNLSLVE 629
QY 653 NVDPKVIELYTRDEKRLSNVFDIETMPREYNYAQVSGRLRKDLLD-LSTNRNQFVDAFG 711
DB 630 NIDFKVIELYSKTEKRLSNVFDIETMPREYNYQNLNGLRKDLLDSDIDHGRDSFIOTLG 689
QY 712 SLMDDLGAVGQTVNVNAVSGVATLFSISIVTGFINIKNPFQGMMLIIVVI-GVLFAYFLT 770
DB 690 DIMQDLGTIGKVVNVNVAAGVFSLFGSIVSGVISFKNPFGGMLLIIVIIAGVVVYLFMT 749
QY 771 KKTIIYETAPIKMIYPEIDKLKEREKSEIAPISSEELERIVLAMHIHQNSHMETKTRK 830
DB 750 RRSRIY-SAPIRMLYPGVERAAQEPGAH---PVSEDIQIRNILMGHMHQFQORAEER 805
QY 831 DPK-----DSILTRAQNMRLKX---SGYSNLK 854
DB 806 EEEVKGKRTLFEVIRDSATSVLRRRGGGYQRLQ 840

RESULT 7
US-09-338-326-16
Sequence 16, Application US/09338326
GENERAL INFORMATION:
APPLICANT: Rose, Timothy M.
APPLICANT: Bosch, Marnix L.
APPLICANT: Strand, Kurt
TITLE OF INVENTION: GLYCOPROTEIN B OF THE RFHV/KSHV
SUBFAMILY OF HERPES VIRUSES
NUMBER OF SEQUENCES: 100
CORRESPONDENCE ADDRESS:
ADDRESSEE: Morrison & Foerster
STREET: 755 Page Mill Road
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304-1018
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/338,326
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/720,229
FILING DATE: 26-SEP-1996
ATTORNEY/AGENT INFORMATION:
NAME: Schiff, J. Michael
REGISTRATION NUMBER: 40,253
REFERENCE/DOCKET NUMBER: 29938-20002.00
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 813-5600
TELEFAX: (415) 494-0792
TELEX: 706141
INFORMATION FOR SEQ ID NO: 16:
SEQUENCE CHARACTERISTICS:

```
;
; LENGTH: 874 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-338-326-16

Query Match      43.0%; Score 1953.5; DB 17; Length 874;
Best Local Similarity 46.1%; Pred. No. 2.2e-174; Indels 75; Gaps 18;
Matches 403; Conservative 147; Mismatches 250;

Qy 11 VLALWLYQVALYSLSAETGVTSPPTATWSTESPLTGHYGDHSHGGRGNENRDS 70
Db 10 VLCLWCV--AALLCQVAQEVVA-----ETTP-----PATH-----RPEVVAE 46

Qy 71 EQNKNIYGSSTPYRVCSAGV-GDVFRTQDTHVCPDASDMVHSEGILLIYKQNIIPFM 129
Db 47 ENPANPE-----LPFRVCGASPTGGEIFRFPLEESCPNTEDKDHIEGIALIYKTNIPVY 101

Qy 130 FRVKYRKVVTTSTVYNGIYSDSLTNOHTFYKSTIEPWETEKMDTIYOCFNSLRNTG 189
Db 102 FNVRYKRIKMTSTIYKGSWSDALTNQHTRSYAVPLVEYQMDHYOCFSAVQVNEGGHV 161

Qy 190 LTYVDRDDINMTVLPQVDGVTDPVKRYGSOPELYLEPGWFSYRRRTTVNCELMDMFA 249
Db 162 NTYVDRDGNWETAFLKPADGLTSSITRYQSOPEYATPRNLLWSYTRTTVNCVEIEMSA 221

Qy 250 RSNPPDFVFTATGDTVEMSGPFWGEGEDDHENKMKHKEKFWFSVINNYKVVDYQNRGT 309
Db 222 RSMKPFVFTVSGDTTIEMSPFLKENGTEPEKILKRPHSLQLLKNKAVATKY---GVGLG 277

Qy 310 K---TRIFLDREBYTILSWKHLKNMSYCPILTAKAFYNGIOTHSYHFVANDITASF 365
Db 278 QADNATRFALFGYLSUSWKATTESSYCDLILWKGFSAIQTQHNSLHFIANDITASF 337

Qy 366 TTSKEDMKFEFTYHCLNEEKIKAEIKYAKVNSHSGYKDLKYFKTDGGLYLWVQPLIQ 425
Db 338 STPLEEAEANFETFKIWNNTQBEIQKKEVEKTHRPNGTKAVKYKTTGNLYIVWQPLIQ 397

Qy 426 NRLLDANKLNN-----ETYSRRSRROAESTTDPMMEMTNGAGGEYSS-ENS 472
Db 398 IDLLDTHAKLYNLNATASPTSTPTSPRRRRRTSS-----VSGGNGNGNSTKEES 450

Qy 473 ITVAQVQAYDNLRIKRNILLEDLSKAWCEQRAALVWNLKINPTSVMSIYNRPVS 532
Db 451 VAASQVQFAYDNLKRSINRVLGELUSRAWCEQYRASLWYELSKINPTSVMSALYGRPV 510

Qy 533 AKRGDVISVNCIVVDQTSVLSHLSRLILSASDEKCFSRPPVTFKPMNDSTIYKQGLGV 592
Db 511 AKLIGDVVSVDCLISVDQKSVFVHKMK-VPGKEDLCYTRPVGVFKFINGSELPAGQLGP 569

Qy 593 NNEILLTTLTLETCENTYFYQAKTDMYIKVNYEHLKTVPLSSITLDTFIALNFTLE 652
Db 570 RNEIVLSTSQVEVCQHSCEHYFQAGNQMYKYDYVYVSTLNLTDIPTLHTMITLNLSLVE 629

Qy 653 NVDEKVELITRDEKRLSNVPDIETMREYNYYAQRVSGLRKLLD-LSTNRQFVDAPG 711
Db 630 NIDPKVELYLSKTEKRLSNVPDIETMREYNYYQNLNGLRKLLDSDSIDHGRDSFIOTLG 699

Qy 712 SLMDDLGAVGQTVNAVSGVATLFSSTIVTGFINPKNPPFGMLMIIVVI-GVLFALYFLT 770
Db 690 DIMQDLGTIGKVVNVASGVFSLFGSIVSGVIFPKNPPFGMLLILVILIAVVVVYLFMT 749

Qy 771 KTKIYETAPIKMYPIDKLKEREGKEIAPISEEBLERIVLAWHIHQONSHMETKTRK 830
Db 750 RRSRIY-SAPIRMLYPGVARAAQPGAH---PVSEDDQIRNLMGMHQFQQRAREEBARR 805

Qy 831 DPK-----DSLTRAQNMRLKX---SGYSNLK 854
Db 806 EEEVKGKRTLFEVIRDSATSVLRRRRGGGGYQLQ 840

RESULT 8
```

```
US-08-804-439-14
; Sequence 14, Application US/08804439
; GENERAL INFORMATION:
; APPLICANT: Rose, Timothy M.
; APPLICANT: Bosch, Marnix L.
; TITLE OF INVENTION: GLYCOPROTEIN B OF THE RPHV/KSHV
; NUMBER OF SEQUENCES: 113
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morrison & Foerster
; STREET: 755 Page Mill Road
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304-1018
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; APPLICATION DATA:
; FILING DATE:
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Schiff, J. Michael
; REGISTRATION NUMBER: 40,253
; REFERENCE/DOCKET NUMBER: 29938-20002.20
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 813-5600
; TELEFAX: (415) 494-0792
; TELEX: 706141
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 808 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-804-439-14
```

```
Query Match      42.9%; Score 1949.5; DB 12; Length 808;
Best Local Similarity 46.6%; Pred. No. 4.6e-174; Indels 33; Gaps 10;
Matches 373; Conservative 159; Mismatches 236;

Qy 71 EQNKNIYGSSTPYRVCSAGVGDVFRFQDTHVCPDASDMVHSEGILLIYKQNIIPFM 130
Db 28 EKUKTQAIQYFVKYRVCSASTTGLFRLDRTCPSTEDKVHKEGILLIYKKNIVPIY 87

Qy 131 RVKRYKVVTTSTVYNGIYSD--SITNQHTFYKSIETPWEKMDTIYOCFNSLRNTG 188
Db 88 KVRRYKKTITSVRIFNGWTRREGVAITNKWELSAVPKYEIDIMDKTYQCHNCMQIEVNGM 147

Qy 189 LTYVDRDDINMTVLPQVDGVTDPVKRYGSOPELYLEPGWFSYRRRTTVNCELMDMF 248
Db 148 LNSYDRDGNKNTVDUKPVDGLTGALTRYISQKPFADPCWLWGTYTRTTVNCIYDMF 207

Qy 249 ARSNPDPDFVFTATGDTVEMSGPFWGEGEDDHENKMKHKEKFWFSVINNYKVVDYQNRGT 308
Db 208 ARSADPTVFTVFTALGDTVEVSPFCVDVNSCPNATDVL--VQIDLNHTVVDYGNRATSOQ 265

Qy 309 GKTRIFLDREBYTILSWKHLKNMSYCPILTAKAFYNGIOTHSYHFVANDITASF 368
Db 266 HKKRIFAHTLDYSVSEAVNKSASVSMVFWKSFQRAIQTEHDLTYHFIANEITAGFSTV 325

Qy 369 KEDMKFEFTYHCLNEEKIKAEIKYAKVNSHSGYKDLKYFKTDGGLYLWVQPLIQ 428
Db 326 KEPLANTSDYNCLMTHINTTLEDKJARVNNHTPNGTAETAYYQTEGMLLVWQPLIAEL 385

Qy 429 LDKANKLNNE-----TYSRRSRROAESTTDPMMEMTNGAGGEYSSSENSITVAQVYA 481
Db 386 EEALEATTSPVTPSAPTSSRSKRAIRSIRDV-----SAG-----SENNVLSQIOYA 434
```

```
QY 482 YDNLIRINNIILEDLSKAWCREQRAALVNNELSKINPTSVMSIYNRPVSAKRIGDVIS 541
Db 435 YDKLRQSINNVLLEALITWCREQVQRTWVYIAKINPTSVMTAIYKRPVSRKALGDVIS 494
QY 542 VNCIVVDQTSVSLHKSLLLSASDEKCFRPPVTFKFMNDSTIYKQGLGVNNELLTTT 601
Db 495 VTECINVDQSSVSIHSLK--TENNDICYSRPPVTFKFMNSQLFKQGLGARNEILLSES 552
QY 602 YLETCQENTYVFOAKTDMYIKNYEHLKTVPLSSITLDTFIALNFTLLENVDFKVEL 661
Db 553 LVENCHQNAETFFAKNETYHFKNYVHVELPUNNISTLDTFLALNLTFIENIDFKAVEL 612
QY 662 YTRDEKRLSNVFDIETMFREYNYAQRVSGRLKDLDDL--TNRNQFVDAFGSLMDDLGA 719
Db 613 YSSGERKLANVFDLETMFREYNYAQSISGLRKD-FDNSQRNRRDRIIQDFSEILADLGS 671
QY 720 VQQTVVNAVSGVATLPSISVITGFINFKNPFGMLMIIVIGLVFAIYFLTKTKIYETA 779
Db 672 IGKIVNVASGAFSLFGGIVTGILNFKNPLGGMFTLLIGAVIILVILLVRRTNMNSQA 731
QY 780 PIKMIYPEIDKLKEREGSEIAPISEEELERIVLAWHIHOONSHMETKTRKDPKDSILTR 839
Db 732 PIRMIYPDVEK-----SKSTVTPMEPETIKOILLGHMNMQOEAYKKKEQRAARPSIFRQ 786
QY 840 -AQNMLRKRSYGYNLKNAESV 859
Db 787 AAEFLKRSYGKQISTEDKI 807

RESULT 9
US-09-301-390-14
; Sequence 14, Application US/09301390
; GENERAL INFORMATION:
; APPLICANT: Rose, Timothy M.
; APPLICANT: Bosch, Marnix L.
; APPLICANT: Strand, Kurt
; TITLE OF INVENTION: GLYCOPROTEIN B OF THE RPHV/KSHV
; TITLE OF INVENTION: SUBFAMILY OF HERPES VIRUSES
; NUMBER OF SEQUENCES: 100
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morrison & Foerster
; STREET: 755 Page Mill Road
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304-1018
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/301.390
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/720,229
; FILING DATE: 26-SEP-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Schief, J. Michael
; REGISTRATION NUMBER: 40,253
; REFERENCE/DOCKET NUMBER: 29938-20002.00
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 813-5600
; TELEFAX: (415) 494-0792
; TELEX: 706141
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 808 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
```

```
MOLECULE TYPE: protein
US-09-301-390-14
Query Match 42.9%; Score 1949.5; DB 17; Length 808;
Best Local Similarity 46.6%; Pred. No. 4.6e-174;
Matches 373; Conservative 159; Mismatches 236; Indels 33; Gaps 10;
QY 71 BONKNYIGSPSTFPYVCSASGVDFRFQDTHVCPDASDMVHSEGLLIYKONIPMF 130
Db 28 EKNTQAIYQEFYKRVCSASTTGLFPFDLDRTCPSTEDKVHREGILLVYKKNIYVIP 87
QY 131 RVRYKRVKVTSTVYNGIYSD--SITNQHTFYKSIEMETEKMDTIYOCFNSLRANTGN 188
Db 88 KVRKYKITTIVRIFNGWTREGVAITNKWELSRAPVKEIDIMDKTYOCHNCMQIEVNGM 147
QY 189 LLTVVDSDDDINMTVFLQPDGVTPDVKRYGSOPELYLEPGHFWGFSYRRRTTNCCLMDMP 248
Db 148 LNSYVDRDGNKNTVDLPVGLTGATITRYISQKPFADPGMLWGTYRTRTTVNCIEIVDMF 207
QY 249 ARSNPPDFFTATGDTVMSPFSGEDDHENKHEKPFVSVNNYKVDYQNRGTVP 308
Db 208 ARSADPTFYFTALGDIIVEVSPFCDDVNSCPNATDVL--VOIDLNHTVVDYGNRATSOQ 265
QY 309 GKTRIFLDREBYTILSWELKLNMSYCPULTLWAPYNGIQTEHSGSYHFVANDITASFPTS 368
Db 266 HKRIFAHTLDYSVSEAVNKSASVCSMVFWKSFQRAIQTEHDLTYHFIAEITAGFSTV 325
QY 369 KEDMKKEFTYHCLNEEIKAEIEKKAIVNSTHSGYKGLDYKFTDGGLYLWQPLIQNRL 428
Db 326 KEPLANFTSYNCLMTHINTTLEDKIARVNNTHTPNGTAEYIQTGEGMILWQPLIAIEL 385
QY 429 LDKNKLNE-----TYSRRSRQAEISTTDPMMWMTGNGAGGEYSSENSITVAQVOYA 481
Db 386 BEAMLEATTSPVTPSAPTSSRSKRARSIRDV-----SAG-----SENNVFUSIQIYA 434
QY 482 YDNLIRINNILEDLSKAWCREQRAALVNNELSKINPTSVMSIYNRPVSAKRIGDVIS 541
Db 435 YDKLRQSINNVLLEALITWCREQVQRTWVYIAKINPTSVMTAIYKRPVSRKALGDVIS 494
QY 542 VNCIVVDQTSVSLHKSLLLSASDEKCFRPPVTFKFMNDSTIYKQGLGVNNELLTTT 601
Db 495 VTECINVDQSSVSIHSLK--TENNDICYSRPPVTFKFMNSQLFKQGLGARNEILLSES 552
QY 602 YLETCQENTYVFOAKTDMYIKNYEHLKTVPLSSITLDTFIALNFTLLENVDFKVEL 661
Db 553 LVENCHQNAETFFAKNETYHFKNYVHVELPUNNISTLDTFLALNLTFIENIDFKAVEL 612
QY 662 YTRDEKRLSNVFDIETMFREYNYAQRVSGRLKDLDDL--TNRNQFVDAFGSLMDDLGA 719
Db 613 YSSGERKLANVFDLETMFREYNYAQSISGLRKD-FDNSQRNRRDRIIQDFSEILADLGS 671
QY 720 VQQTVVNAVSGVATLPSISVITGFINFKNPFGMLMIIVIGLVFAIYFLTKTKIYETA 779
Db 672 IGKIVNVASGAFSLFGGIVTGILNFKNPLGGMFTLLIGAVIILVILLVRRTNMNSQA 731
QY 780 PIKMIYPEIDKLKEREGSEIAPISEEELERIVLAWHIHOONSHMETKTRKDPKDSILTR 839
Db 732 PIRMIYPDVEK-----SKSTVTPMEPETIKOILLGHMNMQOEAYKKKEQRAARPSIFRQ 786
QY 840 -AQNMLRKRSYGYNLKNAESV 859
Db 787 AAEFLKRSYGKQISTEDKI 807

RESULT 10
US-09-338-326-14
; Sequence 14, Application US/09338326
; GENERAL INFORMATION:
; APPLICANT: Rose, Timothy M.
; APPLICANT: Bosch, Marnix L.
; APPLICANT: Strand, Kurt
; TITLE OF INVENTION: GLYCOPROTEIN B OF THE RPHV/KSHV
; TITLE OF INVENTION: SUBFAMILY OF HERPES VIRUSES
```


TELECOMMUNICATION INFORMATION:

TELEPHONE: (415) 813-5600
 TELEFAX: (415) 494-0792
 TELEX: 706141

INFORMATION FOR SEQ ID NO: 94:

SEQUENCE CHARACTERISTICS:

LENGTH: 845 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear

FEATURE:

NAME/KEY: Modified-site

LOCATION: 841

OTHER INFORMATION: /note= "Proline or Leucine

OTHER INFORMATION: depending on codon"

US-09-338-326-94

Query Match 42.6%; Score 1937; DB 17; Length 845;

Best Local Similarity 48.5%; Pred. No. 7.6e-173;

Matches 382; Conservative 135; Mismatches 237; Indels 34; Gaps 10;

Qy 83 FPRVCSAGVGDFRFDHVCPCDASDMVHSEGLLIYKONIIPFMRVKRYKVVTT 142

Db 63 YQFVCSASITGELFRNLEQCPDTKQKHOGELLVYKKNIVPHIFKVRKYRIATSV 122

Qy 143 TYVNGIYSDSITNQHFTFKSIEPWEKMDTIYQCFNSLRNLNTGGLLTYVDRDDINMTV 202

Db 123 TVYRGLTESAITNKYELPRPPLYEISHMDSYQCFSSMKVNVNGVENTFTDRDDVNTV 182

Qy 203 FLOPDGVTDPVKRYGSOPELYLEPGFWGVSRRRTTNCCELMDFARNPPDFVFTAT 262

Db 183 FLQPVGLTNTQRYFSQPIVAEPGFPPIYVRRTVNCCEIVDMIARSABPNYFVTS 242

Qy 263 GDTVMSPPWSEDDHENKHEKPF-VSVINNYKVDYQNGTVPGLKTRIFLDREYT 321

Db 243 GDTVMSPPWSEDDHENKHEKPF-VSVINNYKVDYQNGTVPGLKTRIFLDREYT 302

Qy 322 LSWKHLKMSYCPITLWKAIFYQTEHSGSYHEVANDITASFTTSKEDMKEFNNTYHC 381

Db 303 LSWASESKTAVCPALAKTKTPRSQTTTHDSFHEVANEITATFTAPLTPVANFTDTYSC 362

Qy 382 LNEETKAEIKYAKNTHSKYDGLKFKTDGGLYLVQPLIQNLRLDANKLNNET-- 439

Db 363 LTSINTTLNASKAKLASHVPGVTQVYFHTTGGLYLVQPMASALNTHAQGDSGNPTSS 422

Qy 440 -----YSRSRRQAESETTPMMEGNGAGGEYSSENSITVAQVAYDNLIR 488

Db 423 PPSASPMTTSASRRKRSASTA-----AAGGGGSTDN-LSYQLOFAYDKLRDG 471

Qy 489 INNILEDSKAWCRQHRRAALVWNLKINPTSVMSIYNRPVSASAKRGDVISYNSCIW 548

Db 472 INQVLELSRAWCRQVRDNLWMLKINPTSVMTAIYGRPVSAKFVGDAISVTECLNV 531

Qy 549 DQTSVSLHKSURLLSASDEKCSRPVTFKFMNDSTIYKGLGVNNEILLTTTYLETQCE 608

Db 532 DQSSVNIHKSRL--TNSKDVCVARPLVTFKFLNSSLNFTGOLGARNELIITNNQVETCKD 589

Qy 609 NTEYFQAKDMYIKVNYEHLKTVPLSSITLDTLFIALLNFLENVDKFIYELITRDEKR 668

Db 590 TCEHVFITRNEFLVYKVDYARTINTTIDISTLNTFIALLNLSFIQINDFKALIELYSSAEK 649

Qy 669 L-SNVFDIETPREYNYAQRVSGURLKDLLO-LSTNRNQFVDAFGSLMDDLGAAGVOTVVN 726

Db 650 LASSVFDLETWPREYNYTHRLAGLRDLDDNTIDNKKERFVRDLSEIVADLGGIGKTVVN 709

Qy 727 AVSGVATLFSSIVTGFINFKNPFGLMLMIIVVIGLVFAIYFLTKTKIYETAPTKMIYP 786

Db 710 VASSVVTLCGSLVTGFINFKHPLGLMLMIIVIAIILIFMLSRRTNTIAQAPVMIYP 769

Qy 787 EIDKLKEREGKSEIAPISEBELERVLAMHITHQNSHMETKTRDKPKDSILTRQNMRLK 846

Db 770 DVDR---RAPPSGGAP-TREEIKNLLGWHQLOQBERQAKDCLKKSTPSVFQRTANGLRQ 825

Qy 847 R-SGYSNL 853

Db 826 RLRYKPL 833

RESULT 15

US-09-053-601-3

; Sequence 3, Application US/09053601

; GENERAL INFORMATION:

; APPLICANT: Desrosiers, Ronald C.

; TITLE OF INVENTION: NOVEL HERPESVIRUS

; NUMBER OF SEQUENCES: 6

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Banner & Witcoff, Inc.

; STREET: 28 State Street, 28th Floor

; CITY: Boston

; STATE: Massachusetts

; COUNTRY: USA

; ZIP: 02109

; COMPUTER READABLE FORM:

; MEDIUM TYPE: 3.5" high density diskette

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: WordPerfect 6.1

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/053.601

; FILING DATE: 01-APR-1998

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 60/043,664

; FILING DATE: 11-APR-1997

; ATTORNEY/AGENT INFORMATION:

; NAME: Kathleen M. Williams

; REGISTRATION NUMBER: 34,380

; REFERENCE/DOCKET NUMBER: 10498/74421

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 617-227-7111

; TELEFAX: 617-227-4399

; INFORMATION FOR SEQ ID NO: 3:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 829 amino acids

; TYPE: amino acid

; STRANDEDNESS: single

; TOPOLOGY: unknown

; MOLECULE TYPE: protein

US-09-053-601-3

Query Match 42.4%; Score 1927.5; DB 14; Length 829;

Best Local Similarity 44.2%; Pred. No. 5.8e-172;

Matches 391; Conservative 145; Mismatches 245; Indels 103; Gaps 15;

Qy 11 VLALWLYQVALYSLSTAETGVTSPNPTATWSTESPLTGHYTHDSSHGRCNNENRDS 70

Db 10 LLRAWVY--TIAIGTAVGE-NVTTPKGATTTAKPTP----- 42

Qy 71 EQNKNIYGPST-----FPRVCSAGVGDFRFDHVCPCDASDMVHSEGL 118

Db 43 -----GSTPTFPENPRAEAFKRVCSASATGELFRNLEKTCPTGCTEDKTHQEGIL 94

Qy 119 LIYQNIIPFMRVKRYKVVTTSTVYNGIYSDSITNQHTFYKSIPEWTEKMDTIYQCF 178

Db 95 MVFKKNIIVPHIFKVRKYRVKAVTSVYVGMWTAVTGKQEVIRPVQYEHNMDDTTYQCF 154

Qy 179 NSLRNLNTGGLLTYVDRDDINMTVFLQPDVGTVPDKRYGSOPELYLEPGFWGVSRRRT 238

Db 155 SSMRVNVNGIVNTYDRDFTNQTVFLQPVLEGLTDNIQRYFSQPVLYTTPGMPFGIYVRT 214

Qy 239 TVNCELMDFAARNPPDFPVTTATGDTVMSPFWSG-----DDHENKHEKPFVSVI 292

Db 215 TVNCEIVDMIAARSAEPIYFVTALGDTVEVSPFCHNDSTCSVAEKTENGLGAR-----VL 269

Qy 293 NNYKVDYQNGRTVPLGKTRIFLDREBYTLSWEKHLKNNMSYCPITLWKAIFYQTEHSG 352

Db 270 TNYTMVDFATR--APTETRVFADSGEYTVSWRAEDPKSAVCALTLWKTFPRAIQTTHA 327


```
Query Match 42.2%; Score 1918; DB 26; Length 823;
Best Local Similarity 48.5%; Pred. No. 4.6e-171;
Matches 378; Conservative 134; Mismatches 233; Indels 34; Gaps 10;

QY 83 PPRVCSASGVDFRFRQDTHVCPDASDMVHSEGLLIYKONIIPFMRVRYKRVKVTTS 142
DB 63 YOFVCSASITGELFRNLEQTCPTDKYHQEGILLVKKNIIVPHIFKVRRIYIATSV 122
QY 143 TVYNGIYSDSINOTFFKSTPEWETEKMDTIYQCFSNLSRLNTGNGLLTYVDRODDINMTV 202
DB 123 TVYRGLTSATINKYELRPPVLYEISHMDSYIQCFSKKNVNVNGVENTFTDRDDVNTV 182
QY 203 FLQPDVGVTPDKVRYGSOPELYLEPGWFGSRRRTTNCMLDMFARSNPPDFPVAT 262
DB 183 FLQPEGLTDNIQRYFSQPVIAEPGFWPGVIRVTRTVNCEIVDMIAESAPYNYFVTSL 242
QY 263 GDTVEMSPFWGDEDDHENKHEKPFVSVINNYKVVDYQNRGTVPGLKTRIFLDREY 321
DB 243 GDTVEVSPFCYNSSCSTTPSNKGLSVQVLNHTVWYSDRGTSPTPQNRIFVETGAY 302
QY 322 LSWKHLKNMSYCPULTKAFYNGIOTBSHSGSYHFVANDITASFTTSKEDMKEFNTHVC 381
DB 303 LSWASEKTTAVCPALWKTTPRSLOTTHDSHFHFVANEITATFTAPLTPVANFTDTYSC 362
QY 382 LNEEIKABIEKKYAKVNSTHSGYDLYFKTDGGLYLVWQPLIQNRLLDANKLNNET-- 439
DB 363 LTSDINTLNASKAKLASTHPVNGTVQYFHTTGLYLVWQPMASALNTHAQDGSNPTSS 422
QY 440 -----YRSRRROAESTTDPMMETGNGAGGYSSENSITVAQVAVYNDLRIR 488
DB 423 PPPSASPMTTSASRRKRSSASTA-----AAGGGGSTDN-LSYTLQLOFAYDKLRDG 471
QY 489 INNILEDLKAWCREQHRALVWNLKINPTSVMSMIYNRPVSAKRGDVISVNCIIV 548
DB 472 INQVLELSRAWCREQVNDLMWYELSKINPTSVMTAIGRPVSAKFVGDALSVTECIN 531
QY 549 DQTSVSLHSLRLLSASDEKCFSPRPVTFKPMNDSTIYKGLGVNNEILLTTVLETQOE 608
DB 532 DOSSVNIHKSRL--TNSKDCVAPLVTFKFLNSSNLFTGQLGARNEILLTNQVETCKD 589
QY 609 NTEYFQAKTDMYIYKNEYHLKTVPLSSITLDTFIALNFTLLENVDFKVIETYRDEKR 668
DB 590 TCEHYFTRNETLVYKDYAVLRTINTDITLNTFIALNLSFIQNDKAIELYSSABKR 649
QY 669 L-SNVFDIETMFREYNYAQRVSLGRKDLDD--LSTNRNQFVDAFGLSMDLGAQGVTVN 726
DB 650 LASSVFDLETMFREYNYTHRLAGUREDLNNTIDMKNKERFVRDLSEIVADLGGIGKTVN 709
QY 727 AVSGVATLFSSTVTFINFIKNPFGGMLMIWIGVLFALYFLTKTKIYETAPIKMIYP 786
DB 710 VASSVVTLCGLSVTGFINFIKHPLGGMLMIIVIAILLIIFMLSRNTTIAQAPVKMIYP 769
QY 787 EIDKLEREGKSEIAPISEELERIVLAWHIHQONSHMETKTRKDPKDSILTRAQNMLR 845
DB 770 DVDR---RAPPSGGAP-TREIEKNILLGWHQLQOERQKADDLKKS-TPSVFQRTANGLR 823

RESULT 18
US-10-055-364-42
; Sequence 42, Application US/10055364
; GENERAL INFORMATION:
; APPLICANT: Patience, Clive
; TITLE OF INVENTION: Gamma Herpes Virus DNA and Methods of Use
; FILE REFERENCE: 61750-379
; CURRENT APPLICATION NUMBER: US/10/055,364
; CURRENT FILING DATE: 2002-01-23
; PRIOR APPLICATION NUMBER: US/09/612,204
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: US/60/142,736
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: US/60/168,532
; PRIOR FILING DATE: 1999-12-02

; NUMBER OF SEQ ID NOS: 55
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 42
; LENGTH: 792
; TYPE: PRT
; ORGANISM: Herpesvirus saimiri
US-10-055-364-42

Query Match 42.1%; Score 1916; DB 26; Length 792;
Best Local Similarity 46.9%; Pred. No. 6.6e-171;
Matches 366; Conservative 155; Mismatches 228; Indels 32; Gaps 9;

QY 71 EQNKNIYGSPTPYRVCSASGVDFRFRQDTHVCPDASDMVHSEGLLIYKONIIPMF 130
DB 28 EKNKTQAIYQBYFYKRVCSASTTGCFLFRDLDRDTCPTEDKVKHEGILLVKKNIIVPIY 87
QY 131 RVRYKRVKVTTSITVYNGIYSD--SITNQHTFYKIEPWEKTEKMDTIYQCFSNLSRLNTG 188
DB 88 KVRKYKKTTSVRIFNGWTRGVATNKNWLSRAVPKYEIDIMDKTYQCHNCMIENVGM 147
QY 189 LLTYVDRDDINMTVFLQPDVGVTPDKVRYGSOPELYLEPGWFGSRRRTTNCMLDMF 248
DB 148 LNSYDRDGNKKTVDLPVGLTGCAITRYISQPKVADPGLWGTYRTRITVNCIIVDMF 207
QY 249 ARSNPPDFFTATGDTVEMSPFWGDEDDHENKHEKPFVSVINNYKVVDYQNRGTVP 308
DB 208 ARSADPYTFVTALGDIVESPPFCVDVNSCPNATDVLS--VQIDLNHTVVDYGNRATSSQ 265
QY 309 GKTRIFLDREYTLUSWEKHLKNMSYCPULTKAFYNGIOTBSHSGSYHFVANDITASFTTS 368
DB 266 HKKRIFAHTLDYSVSWEAVNKSASVCSVMVFWKSPQRAIOTEDHDLTYHFIAEITAGFSTV 325
QY 369 KEDMKEFNTHVCNLEEBIEKAEIEKKYAKVNSTHSGYDLYFKTDGGLYLVWQPLIQNR 428
DB 326 KEPLANTSDYNCLMTHINTTLEDKIAKVNHTTPNGTAETIYQTEGGMLVWQPLIAEL 385
QY 429 LDKANKLNNE-----TYRSRRROAESTTDPMMETGNGAGGYSSENSITVAQVQYA 481
DB 386 EAMLEATTSPVTPSAFTSSRSKRAIRSIRDV-----SAG-----SENNVFLSQIOYA 434
QY 482 YDNLRIRINILEDLKAWCREQHRALVWNLKINPTSVMSMIYNRPVSAKRIGDVIS 541
DB 435 YDKLRQSIINNLEELAITWCREQVQRQTMVWYEAIAKINPTSVMTAIGKPKVSRKALGDVIS 494
QY 542 VSNICIVVDQTSVLSHSLRLLSASDEKCFSPRPVTFKPMNDSTIYKGLGVNNEILLTTT 601
DB 495 VTECINVDQSSVSLHSLK--TENNDICYSRPPVTFKPVNSQLFKGOLGARNEILLSES 552
QY 602 YLETQCENTEYFQAKTDMYIYKNEYHLKTVPLSSITLDTFIALNFTLLENVDFKVIEL 661
DB 553 LVENCHQNAETFTTAKNETYHFKNYVHVELPVNNISTLOTFLALNLTFIENIDFKAVEL 612
QY 662 YTRDEKLSNVFDIETMFREYNYAQRVSLGRKDLDDLS--TNRNQFVDAFGLSMDLGA 719
DB 613 YSSGERKLANVFDLETMFREYNYAQRVSLGRK--FDNSQRNRRDRIIQDPSEILADLGS 671
QY 720 VQGVTVNAVSGVATLFSSTVTFINFIKNPFGGMLMIWIGVLFALYFLTKTKIYETA 779
DB 672 IGKVINNVASGAFSLFGGIVTGILNFIKNPLGGMFTFLIGAVIILVILLVRRNNMSQA 731
QY 780 PIKMIYPEIDKLEREGKSEIAPISEELERIVLAWHIHQONSHMETKTRKDPKDSILTR 839
DB 732 PIRMIYFDVEK-----SKSVTPMEPETIKQILLGMHMQOEAYKKKEQRAARPSIFRQ 786
QY 840 A 840
DB 787 A 787

RESULT 19
US-10-055-364-38
; Sequence 38, Application US/10055364
; GENERAL INFORMATION:
```



```

QY 596 ILLTTYLETCEYFQAKTDMYIYKNYHLKTVPLSSITTLDTFIALNFTLLENVD 655
D 596 ILLTTYLETCEYFQAKTDMYIYKNYHLKTVPLSSITTLDTFIALNFTLLENVD 655
Db 546 ILLSESLVENCHQNAEHFFTAQNETVHPKYLHVETPLTNISTLDTFLALNLTIENID 605
QY 656 FKVIELYTRDEKRLSNVFDIETMFREYNYAQRVSLGRKDLDDLDS--TNRNOFVDAFSL 713
D 656 FKVIELYTRDEKRLSNVFDIETMFREYNYAQRVSLGRKDLDDLDS--TNRNOFVDAFSL 713
Db 606 FKAVELYSSEGERKLANVFLDTEMFREYNYAQSISGLRKD-FDNRQNRDRRIIQQFSEI 664
QY 714 MDDLGAQGVTVNAVSGVATLFSSIVTGFINKPFGGMLMIIVIGVLFALYELTKKT 773
D 714 MDDLGAQGVTVNAVSGVATLFSSIVTGFINKPFGGMLMIIVIGVLFALYELTKKT 773
Db 665 LADLGSIGKVIYNIASSAFSLFGVITGLNFKPLGGLMFLVGAIIILVILVVRT 724
QY 774 KIYETAPIKMIYPEIDKLEREGKSEIAPISEEELERIVLAMHIIHQONSHMETKTRKDPK 833
D 774 KIYETAPIKMIYPEIDKLEREGKSEIAPISEEELERIVLAMHIIHQONSHMETKTRKDPK 833
Db 725 NMSQAPIRMIVPDIEK-----SRSSVTPTEPEVIKQILLGHNHQQOBEYKKEEHKASQ 779
QY 834 DSILTRA 840
D 780 PSFLKRA 786

RESULT 21
US-10-055-364-44
; Sequence 44, Application US/10055364
; GENERAL INFORMATION:
; APPLICANT: Patience, Clive
; TITLE OF INVENTION: Gamma Herpes Virus DNA and Methods of Use
; FILE REFERENCE: 61750-379
; CURRENT APPLICATION NUMBER: US/10/055,364
; CURRENT FILING DATE: 2002-01-23
; PRIOR APPLICATION NUMBER: US/09/612,204
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: US/60/142,736
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: US/60/168,532
; PRIOR FILING DATE: 1999-12-02
; NUMBER OF SEQ ID NOS: 55
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 44
; LENGTH: 818
; TYPE: PRT
; ORGANISM: Equine herpesvirus 5
US-10-055-364-44

Query Match
Best Local Similarity 41.2%; Score 1871.5; DB 26; Length 818;
Matches 368; Conservative 149; Mismatches 223; Indels 57; Gaps 13;

QY 85 YRVCSAGSGDVFRFOTDHCPCDASDMVHSEGIILYKQNIIPFMRVRYKRVVTTSTV 144
D 85 YRVCSAGSGDVFRFOTDHCPCDASDMVHSEGIILYKQNIIPFMRVRYKRVVTTSTV 144
Db 56 FVACGASPGGEIIFRPLENCENTEDKEHVEGILLIYKTNIVPIYFNVKRVKLTSTTI 115
QY 145 YNGIVSDSTTQHTYKSIETPHEKMDTYIQCPSNLRNTGNNLTYYVDRDDINMTYVL 204
D 145 YNGIVSDSTTQHTYKSIETPHEKMDTYIQCPSNLRNTGNNLTYYVDRDDINMTYVL 204
Db 116 YKGWSQDAITNQYTSFAMPLPEARLDVNYECYNGIQTENGHLTTVVDRDGYNESVRL 175
QY 205 QVGVGTPDVKYGSOPELYLEPGHFWGSYRRRTTVNCELMDMFARSNPPDFVATGCD 264
D 205 QVGVGTPDVKYGSOPELYLEPGHFWGSYRRRTTVNCELMDMFARSNPPDFVATGCD 264
Db 176 VPADGLTSSIRYHSGOPELYVFPRLNLSYTRTTTWNCEVIDMTARSHKPEYFVTASGD 235
QY 265 TVEMSGPFWSGEDHENKHEKWFVSNVYKQVQNRGTVPGLCK---TRIFLDREY 320
D 265 TVEMSGPFWSGEDHENKHEKWFVSNVYKQVQNRGTVPGLCK---TRIFLDREY 320
Db 236 SIETSPFYF-----NASRRVP--VQVLYNYSVTY-----GVGLSGENVTRFFATLNDP 283
QY 321 TLSWEKHLKMSYCPPLTLKAFYNGIQTESHSYHFVANDITASFTTS-KEDMKEFNTTY 379
D 321 TLSWEKHLKMSYCPPLTLKAFYNGIQTESHSYHFVANDITASFTTS-KEDMKEFNTTY 379
Db 284 SISWKAATENSSYCPPLVWKGPSPAIQTKEKSYHFIADAVTASFTTPTLTDSTSYFNTTY 343
QY 380 HCLNEIEIKAEIKKAKVNSTHYSKYGDLKYFTDGLGLVWQPLIQNLRLDANKLNK-- 437
D 380 HCLNEIEIKAEIKKAKVNSTHYSKYGDLKYFTDGLGLVWQPLIQNLRLDANKLNK-- 437
Db 344 QCAWQDIEGEOIKRDPVSKTHARNGSVQIYKTSNGLYVWQPLVQLDLAAHAKTINST 403

```

```

QY 438 -----ETYSRRSRROAESTTDPMMEMTNGAGGEYSSENSITVAQVQYAYD 483
D 438 -----ETYSRRSRROAESTTDPMMEMTNGAGGEYSSENSITVAQVQYAYD 483
Db 404 DNSSTPTTAPNTTTSRRRRRTGTNTA-----TNNSSNNSSMEENLATSQVQPAYD 457
QY 484 NLRIRINILBDLSKAMCREOHRALVWNLKSKINPTSVMSMIYNRPVSAKRIGDVISVS 543
D 484 NLRIRINILBDLSKAMCREOHRALVWNLKSKINPTSVMSMIYNRPVSAKRIGDVISVS 543
Db 458 QLRKSINRVLBQLSRVCMQNYRASLMWYELSKINPTSVMSAIYGRPVSAKLVDGVQVQS 517
QY 544 NCIVVDOTSVLSHKLRLLSASDEKCFSPRPVPTFKMNDSTIYKGLGNNNEILLTTTLYL 603
D 544 NCIVVDOTSVLSHKLRLLSASDEKCFSPRPVPTFKMNDSTIYKGLGNNNEILLTTTLYL 603
Db 518 DCITVQBSVVRHNL-R-VPGSKOLCTYRPPVVGFKFNGSLFVQGLGARNEILLSTNLV 576
QY 604 ETCENTEYFQAKTDMYIYKNYHLKTVPLSSITTLDTFIALNFTLLENVDFKVIELYT 663
D 604 ETCENTEYFQAKTDMYIYKNYHLKTVPLSSITTLDTFIALNFTLLENVDFKVIELYT 663
Db 577 EVQCHSCEHYFQGNHIYKYKNEYVSTMTNLTDVPTLHTMTLNLSLVENVDFQVQLYS 636
QY 664 RDEKRLSNVFDIETMFREYNYAQRVSLGRKDLDDL--LSTNRNQFVDAFGLMDDLGAQVQ 722
D 664 RDEKRLSNVFDIETMFREYNYAQRVSLGRKDLDDL--LSTNRNQFVDAFGLMDDLGAQVQ 722
Db 637 QKEKRLSNVFDIETMFREYNYTQNLKRLKDLDDSIHGRDSDFIQFLGDLVQDLVPVGD 696
QY 723 TVNAVSGVATLFSSIVTGFINKPFGGMLMIIVIGVLFALY-FLTCKTKIYETAPI 781
D 723 TVNAVSGVATLFSSIVTGFINKPFGGMLMIIVIGVLFALY-FLTCKTKIYETAPI 781
Db 697 VIVNAVSGVSLFSGISVSGVIFSLKPLGALTITLALVGGIIVLYLFTIRRTVYQ-API 755
QY 782 KMIYPEIDKLEREGKSEIAPISEEELERIVLAMHIIHQONSHMETKTRKDPKDSILTRAQ 841
D 782 KMIYPEIDKLEREGKSEIAPISEEELERIVLAMHIIHQONSHMETKTRKDPKDSILTRAQ 841
Db 756 RMLYPEVD---RAPQNVQPIPEDQVRSILLAMHQFOOOOQOOOQOOO-----E 801
QY 842 NMLKRSYGSLKNAES 858
D 842 NMLKRSYGSLKNAES 858
Db 802 EHTQRRSIFDTIRESTS 818

RESULT 22
US-08-360-107-113
; Sequence 113, Application US/08360107
; GENERAL INFORMATION:
; APPLICANT: Bolognesi, Dani P.
; APPLICANT: Matthews, Thomas J.
; APPLICANT: Wild, Carl T.
; APPLICANT: Barney, Shawn O.
; APPLICANT: Lambert, Dennis M.
; APPLICANT: Petteway, Stephen R.
; APPLICANT: Langlois, Alphonse J.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR INHIBITION
; TITLE OF INVENTION: OF MEMBRANE FUSION-ASSOCIATED EVENTS, INCLUDING HIV
; NUMBER OF SEQUENCES: 137
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/360,107
; FILING DATE: 20-DEC-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Coruzzi, Laura A.
; REGISTRATION NUMBER: 30,742
; REFERENCE/DOCKET NUMBER: 7872-013
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-9741/8864
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 113:

```

```
; SEQUENCE CHARACTERISTICS:
; LENGTH: 857 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
US-08-360-107-113

Query Match      40.7%; Score 1848.5; DB 7; Length 857;
Best Local Similarity 46.1%; Pred. No. 1.9e-164;
Matches 368; Conservative 148; Mismatches 241; Indels 41; Gaps 12;

Qy 81 STEFVYVCSAGSGVDVFRFOTDHCPCD-ASDMVHSEGLLIYKONIIPFMRVKYKVV 139
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
44 TSPFVCELSHGDLFRFSDIOCPFGTRENHTEGLLMVFKDNIIPYSFKVRSYTKIV 103
Qy 140 TTSVYNGIYSDSITNOHTFYKSIPEWETEKMDTIYOCFNSLRNLNTGGNLLTYVDRDDIN 199
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
104 TNLIIYNGVADSVYTNRHEEKFSVDSYETDQMDTIYOCYNAVWKTDGLTRVYVDRDGVN 163
Qy 200 MTVPLOPQVGVTPDVVKYSGOPELYLEPGFWGYSRRRTTVNCBLMDMFARSPDPFFV 259
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
164 ITVNLKPTGGLANGVRRYASQTELYDAPGWLITWYTRTTRTVNCLITDMMAKSNPDPFFV 223
Qy 260 TATGDTVEMSPFWSGEDDHENKMKHPWFVSVINNYKVVDYQNRGTVPGLKTRIFLDREE 319
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
224 TTTGQTVEMSPFYDGK--NKETPHERADSFHVRTNYKIVDYNRGTPQGERAFDLKGT 281
Qy 320 YTLSEKHLKN--MSYCPLTLWKAFYNGIQTEHSGSVHFVANDITASFTTTSKEDMKFNNT 378
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
282 YTLNWK--LENRTAYCPLQHWQTFDSTIATETGKSIHFVDTGTSFPVNTTVGIELPDA 339
Qy 379 YHCLNEIKAEIKKAKVNSTHSGYD-LKYFTDGGLYLVNQPLIQNLRLDANK--KL 435
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
340 FKCIIEQVNTKMHKEVYEAQVDRYTKGQEAITYFITSGGLLAWLPLTPRSLATVKNLTEL 399
Qy 436 NNETYS-----RRSRQAESTTDPMMEMTCNGAGGEYSSENSI 473
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
400 TPTTSSPPSPSPAPSAARGSTPAVLRRRRRDAGNATTP---VPPTAPGKSLGTLNPN 456
Qy 474 TVAQVQAYDNLRIIRINNILEDLSKAWCREQRAALVMNELSKINPTSVMSIYNRPVSA 533
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
457 ATVQIQPAYDSLRLQINRMGLDLARANCLQKQKQNMVRLTKINPTTVMSIYKAVAA 516
Qy 534 KRIGDVISVNCIVVDQTSVLSLHKSLLSASDEKCFSRPVPVTKFNDSTIYKQLGVN 593
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
517 KRLGDVISVSCVPVQATVTLRSMR--VPGSETMCSYRPLVSPSFINDTKYEGQLGTD 575
Qy 594 NEILLTTLTFTCOENTYEFQAKTDMVYKKNVEHLKATVPLSSITTLDTFALNFTLEN 653
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
576 NEIFLTKMTVEVQATSOYIFQSGNEIHVNDYHFKTIELDGIATQTFISLNTSLIEN 635
Qy 654 VDFKVIETLTXREKRLSNVEDIETMFREYNYYAQRVSGLRKDLID--LSTNRNPFVDAFGS 712
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
636 IDFASLELYSRDEQRASNVFDEGIFREYNFQAQNIAGLRKOLDNAVSNRQNFVDGLGE 695
Qy 713 LMDDLGAVGTQVNAVSGVATLFSIVTGFINFKNPFQGMMLIIVVIGLVFAIFYLTKK 772
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
696 LMDSLGVSIGSITNLVSTVGGFLSFLVSGFISFPKPNPFGMLILVAVGVVILVSLTR 755
Qy 773 TKIYETAPIKMIYPEIDKLKERECKSE--LAPISSEELERIVLAMIHOONSHMETKR 829
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
756 TRWMSQFPVQMLYPGIDELAAQHASGEGPGINPISKTELQAIMLA--LHEQNEQKRAAQ 813
Qy 830 KDPKDSILTRAQNMLRKR 847
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
814 RAAGPSVASRALQAADR 831
```

RESULT 23

US-08-804-439-18

; Sequence 18; Application US/08804439

; GENERAL INFORMATION:

```
; APPLICANT: Rose, Timothy M.
; APPLICANT: Bosch, Marnix L.
; APPLICANT: Strand, Kurt
; TITLE OF INVENTION: GLYCOPROTEIN B OF THE RFHV/KSHV
; TITLE OF INVENTION: SUBFAMILY OF HERPES VIRUSES
; NUMBER OF SEQUENCES: 113
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morrison & Foerster
; STREET: 755 Page Mill Road
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304-1018
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/804,439
; FILING DATE:
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Schiff, J. Michael
; REGISTRATION NUMBER: 40,253
; REFERENCE/DOCKET NUMBER: 29938-20002.20
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 813-5600
; TELEFAX: (415) 494-0792
; TELEX: 706141
; INFORMATION FOR SEQ ID NO: 18:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 857 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-804-439-18
```

Query Match 40.7%; Score 1848.5; DB 12; Length 857;

Best Local Similarity 46.1%; Pred. No. 1.9e-164;

Matches 368; Conservative 148; Mismatches 241; Indels 41; Gaps 12;

```
Qy 81 STEFVYVCSAGSGVDVFRFOTDHCPCD-ASDMVHSEGLLIYKONIIPFMRVKYKVV 139
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
44 TSPFVCELSHGDLFRFSDIOCPFGTRENHTEGLLMVFKDNIIPYSFKVRSYTKIV 103
Qy 140 TTSVYNGIYSDSITNOHTFYKSIPEWETEKMDTIYOCFNSLRNLNTGGNLLTYVDRDDIN 199
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
104 TNLIIYNGVADSVYTNRHEEKFSVDSYETDQMDTIYOCYNAVWKTDGLTRVYVDRDGVN 163
Qy 200 MTVPLOPQVGVTPDVVKYSGOPELYLEPGFWGYSRRRTTVNCBLMDMFARSPDPFFV 259
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
164 ITVNLKPTGGLANGVRRYASQTELYDAPGWLITWYTRTTRTVNCLITDMMAKSNPDPFFV 223
Qy 260 TATGDTVEMSPFWSGEDDHENKMKHPWFVSVINNYKVVDYQNRGTVPGLKTRIFLDREE 319
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
224 TTTGQTVEMSPFYDGK--NKETPHERADSFHVRTNYKIVDYNRGTPQGERAFDLKGT 281
Qy 320 YTLSEKHLKN--MSYCPLTLWKAFYNGIQTEHSGSVHFVANDITASFTTTSKEDMKFNNT 378
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
282 YTLNWK--LENRTAYCPLQHWQTFDSTIATETGKSIHFVDTGTSFPVNTTVGIELPDA 339
Qy 379 YHCLNEIKAEIKKAKVNSTHSGYD-LKYFTDGGLYLVNQPLIQNLRLDANK--KL 435
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
340 FKCIIEQVNTKMHKEVYEAQVDRYTKGQEAITYFITSGGLLAWLPLTPRSLATVKNLTEL 399
Qy 436 NNETYS-----RRSRQAESTTDPMMEMTCNGAGGEYSSENSI 473
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
400 TPTTSSPPSPSPAPSAARGSTPAVLRRRRRDAGNATTP---VPPTAPGKSLGTLNPN 456
Qy 474 TVAQVQAYDNLRIIRINNILEDLSKAWCREQRAALVMNELSKINPTSVMSIYNRPVSA 533
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
457 ATVQIQPAYDSLRLQINRMGLDLARANCLQKQKQNMVRLTKINPTTVMSIYKAVAA 516
Qy 534 KRIGDVISVNCIVVDQTSVLSLHKSLLSASDEKCFSRPVPVTKFNDSTIYKQLGVN 593
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
517 KRLGDVISVSCVPVQATVTLRSMR--VPGSETMCSYRPLVSPSFINDTKYEGQLGTD 575
Qy 594 NEILLTTLTFTCOENTYEFQAKTDMVYKKNVEHLKATVPLSSITTLDTFALNFTLEN 653
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
576 NEIFLTKMTVEVQATSOYIFQSGNEIHVNDYHFKTIELDGIATQTFISLNTSLIEN 635
Qy 654 VDFKVIETLTXREKRLSNVEDIETMFREYNYYAQRVSGLRKDLID--LSTNRNPFVDAFGS 712
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
636 IDFASLELYSRDEQRASNVFDEGIFREYNFQAQNIAGLRKOLDNAVSNRQNFVDGLGE 695
Qy 713 LMDDLGAVGTQVNAVSGVATLFSIVTGFINFKNPFQGMMLIIVVIGLVFAIFYLTKK 772
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
696 LMDSLGVSIGSITNLVSTVGGFLSFLVSGFISFPKPNPFGMLILVAVGVVILVSLTR 755
Qy 773 TKIYETAPIKMIYPEIDKLKERECKSE--LAPISSEELERIVLAMIHOONSHMETKR 829
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
756 TRWMSQFPVQMLYPGIDELAAQHASGEGPGINPISKTELQAIMLA--LHEQNEQKRAAQ 813
Qy 830 KDPKDSILTRAQNMLRKR 847
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
814 RAAGPSVASRALQAADR 831
```

Db 457 ATVOIQAYDSURRQINRMGLDLARAWCLEQKRONVRLTKINPTTVMSSYKAVAA 516
QY 534 KRIGDVISVNCIVVDQTSVLSHLKSLRLSASDEKCFGRPPVTFKFMNDSTYKQGLGVN 593
Db 517 KRLGDVISVQCVPVNOATVTLRKSMR-VPGSETWCYSRPLVSFINDTKTYEGQLGTD 575
QY 594 NEILLTYYTLETCQNTYFQAKTDMYIKNYEHLKTVPLSSITLDTFIALNFTLLEN 653
Db 576 NEIFLTKMTEVCOATSOYFQSGNEIHVYNDYHFKTIELDGIATLQTFISLNTSLTEN 635
QY 654 VDFKVIELYTRDEKLSNVFDETFREYNYAQRVSGRLKDLDD-LSTNRNOFVDAGS 712
Db 636 IDFASLELYSRDEQKASNVFDEGIFREYNFOAQNIAGRLKDLDDNAVSGNRNQFVDGLGE 695
QY 713 LMDLGLAVGQTVVNAVSGVATLFSISIVTGFINFKNPPFGGMLMIIVIGVLFAYFLTKK 772
Db 696 LMDSLGSGVQSITNLVSTVGGFLSSLVSGFTSFFKNPFGGMLILVLVAGVILVLSLTR 755
QY 773 TKIYETAPIKMIYPIIDKLKEREKSE---IAPISEELERIVLAMHIHQNSHMETKTR 829
Db 756 TRQMSQQPVMQLYPCIDELAQHASGEGPGINPIKTELQAIMLA--LHEQNOEQKRAAQ 813
QY 830 KDPKDSILTRAQNMLRKR 847
Db 814 RAAGPSVASRALQAARDR 831

RESULT 24

US-09-301-390-18
; Sequence 18, Application US/09301390
; GENERAL INFORMATION:
; APPLICANT: Rose, Timothy M.
; APPLICANT: Bosch, Marnix L.
; APPLICANT: Strand, Kurt
; TITLE OF INVENTION: GLYCOPROTEIN B OF THE RFHV/KSHV
; TITLE OF INVENTION: SUBFAMILY OF HERPES VIRUSES
; NUMBER OF SEQUENCES: 100
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morrison & Foerster
; STREET: 755 Page Mill Road
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304-1018
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/301,390
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/720,229
; FILING DATE: 26-SEP-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Schiff, J. Michael
; REGISTRATION NUMBER: 40,253
; REFERENCE/DOCKET NUMBER: 29938-20002.00
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 813-5600
; TELEFAX: (415) 494-0792
; TELEX: 706141
; INFORMATION FOR SEQ ID NO: 18:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 857 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-301-390-18

Query Match 40.7%; Score 1848.5; DB 17; Length 857;
Best Local Similarity 46.1%; Pred. No. 1.9e-164;
Matches 368; Conservative 148; Mismatches 241; Indels 41; Gaps 12;
QY 81 STFPYRUCSAGVGDVFRFQTDHVCDD-ASDMVISEGILLIYKONIIPFMRVKRYKVV 139
Db 44 TSFPRVCELSHGDLFRFSSDIQCPSTGRENTEGLLMVFKONIIPYSFKVRSYTKIV 103
QY 140 TTSVYNGIYSDSTINOHFTFYKTEPWEKMDTIYQCFNSLRNLTGNGLLTYVDRDDIN 199
Db 104 TNLIIYNGVADSVNREHEKFSVDSYETDMDIYQCYNAVKTGDLTRVYVDRDGVN 163
QY 200 MTVLQPVGDVTPDKRYGSOPELYLEPGWFMGYSRRRTTVNCELMDMFARSPDFPV 259
Db 164 ITVNLKPTGGLANGVRRYASQTELYDAPGMLIWTYRTRTTVNCILITDMAKSNPDDFV 223
QY 260 TATGDTVEMSPFWSGEDDHENKHEKWFVSVINNYKVVDYQNRGTVPGLKTRIFLDREE 319
Db 224 TTTGQTVEMSPFYDGK--NKETPHERADSFHVRTNYKIVDYDNRGTNPQGERAFLDKGT 281
QY 320 YTLSEWHLKN-MSYCPCLTLWKAFYNGIQTEHSGSYHFVANDITASFTTSKEDMKFNTT 378
Db 282 YTLSEW--LENRTAYCPLQHWQTFDSTATETGSIHFVTDGTSFVNTTVGIELPDA 339
QY 379 YHCLNEEIKAEIEKKYAKVNSTHSGYGD-LKYFKTDGGLYLVWQPLIQNRLLDARN--KL 435
Db 340 PKCIEEQVNTMHEKYEAVQDRYTKGOEAIYFITSGLLLAWLPLTPRSLATVKNLTTEL 399
QY 436 NNETY-----RRRRQAESTTDPMMETGNGAGEYSSNSI 473
Db 400 TPTTSSPPSPSPAPSAARGSTPAVLRRRRRRDAGNATP---VPPTAPGKSLGTLLNP 456
QY 474 TPAOVAYDNLRIRINNILEDLSKACROHRAALVWNLKINPTSVMSIYNRPVSA 533
Db 457 ATVOIQAYDSLRQINRMGLDLARAWCLEQKRONVRLTKINPTTVMSSYKAVAA 516
QY 534 KRIGDVISVNCIVVDQTSVLSHLKSLRLSASDEKCFGRPPVTFKFMNDSTYKQGLGVN 593
Db 517 KRLGDVISVQCVPVNOATVTLRKSMR-VPGSETWCYSRPLVSFINDTKTYEGQLGTD 575
QY 594 NEILLTYYTLETCQNTYFQAKTDMYIKNYEHLKTVPLSSITLDTFIALNFTLLEN 653
Db 576 NEIFLTKMTEVCOATSOYFQSGNEIHVYNDYHFKTIELDGIATLQTFISLNTSLTEN 635
QY 654 VDFKVIELYTRDEKLSNVFDETFREYNYAQRVSGRLKDLDD-LSTNRNOFVDAGS 712
Db 636 IDFASLELYSRDEQKASNVFDEGIFREYNFOAQNIAGRLKDLDDNAVSGNRNQFVDGLGE 695
QY 713 LMDLGLAVGQTVVNAVSGVATLFSISIVTGFINFKNPPFGGMLMIIVIGVLFAYFLTKK 772
Db 696 LMDSLGSGVQSITNLVSTVGGFLSSLVSGFTSFFKNPFGGMLILVLVAGVILVLSLTR 755
QY 773 TKIYETAPIKMIYPIIDKLKEREKSE---IAPISEELERIVLAMHIHQNSHMETKTR 829
Db 756 TRQMSQQPVMQLYPCIDELAQHASGEGPGINPIKTELQAIMLA--LHEQNOEQKRAAQ 813
QY 830 KDPKDSILTRAQNMLRKR 847
Db 814 RAAGPSVASRALQAARDR 831

RESULT 25

US-09-338-326-18
; Sequence 18, Application US/09338326
; GENERAL INFORMATION:
; APPLICANT: Rose, Timothy M.
; APPLICANT: Bosch, Marnix L.
; APPLICANT: Strand, Kurt
; TITLE OF INVENTION: GLYCOPROTEIN B OF THE RFHV/KSHV
; TITLE OF INVENTION: SUBFAMILY OF HERPES VIRUSES
; NUMBER OF SEQUENCES: 100
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morrison & Foerster

```
STREET: 755 Page Mill Road
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304-1018
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/338,326
FILING DATE:
CLASSIFICATION:
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 08/720,229
FILING DATE: 26-SEP-1996
ATTORNEY/AGENT INFORMATION:
NAME: Schiff, J. Michael
REGISTRATION NUMBER: 40,253
REFERENCE/DOCKET NUMBER: 29938-20002.00
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 813-5600
TELEFAX: (415) 494-0792
TELEX: 706141
INFORMATION FOR SEQ ID NO: 18:
SEQUENCE CHARACTERISTICS:
LENGTH: 857 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-338-326-18

Query Match 40.7%; Score 1848.5; DB 17; Length 857;
Best Local Similarity 46.1%; Pred. No. 1.9e-164;
Matches 368; Conservative 148; Mismatches 241; Indels 41; Gaps 12;

QY 81 STFPYRVCSAGVDVFRFQTDHVCDD-ASDMVHSEGLLIYKQNIIPFMRVKRYKV 139
DB 44 TSPFPRVCELSHGDLFRFSDIOCPFGTRENHTEGLLMVFKDNIIPYSFKVRSYTKIV 103
QY 140 TTSTVYNGIYSDSTINQHTFYKSIPEWETEKMDTIYOCFNSLRNTGNNLLTYVDRDIN 199
DB 104 TNLIYNGWYADSVTNHREKFSVDSYETQMDTIYOCYNAVXKTKDGLTRVYVDRDGVN 163
QY 200 MTVFLQPDVGTDPVKRYGQPELYLEPGFWGYSRRRTTVNCELMDMFARSNPPDFV 259
DB 164 ITVNLKPTGGLANGVRRYASQTELYDAPGMLIWTYRTTIVNCLITDMMAKSNSPDDFFV 223
QY 260 TATGDTVEMSPFWSGEGDDHKNMHEKFPVSVINNYKVVDYQNGRTVPLGKTRIFLDREE 319
DB 224 TTTGQTVMSPFVDGK--NKETTFHERADSFHVRTNYKIVDYDNRGTNPQGERAFLDKGT 281
QY 320 YTLNWK--MSYCPGLTWKAFYNGIOTHEHSGSYHFVANDITASFTTSKEDMKERNNT 378
DB 282 YTLNWK--LENRTAYCPLQHWQTFDSTIATETGKSIHFVDEGTSFVNTTVGIELPDA 339
QY 379 YHCLNEEIKAEIKKAKVNSTHSGYD-LKYFKTDGGLYLVQPLQNRLLDAKN--KL 435
DB 340 FKCIIEQVNTMHEKYEAVQDRTYTKQEAITYFITSGGLLLAWLPLTPRSLATVKNLT 399
QY 436 NNETY-----RRSROAESTTTPMMEMTGNGAGGEYSNSI 473
DB 400 TTPTSSPPSPSPAPSAARGSTPAALRRRRRRDAGNATTP--VPPTAPGKSLGLTNLP 456
QY 474 TVAQVQAYNLRIRINILDLKAWCRQHRALVWNLKINPTSVMSIYNRPVSA 533
DB 457 ATVOIQFAYSLRQINRMGLDLARAWCLQKRWVLRRLTKINPTTVNSSIYKAVAA 516
QY 534 KRIGDVISVNCIVVQDTSVLSLHKSRLLSASDEKCFSPRPPTFKFMDNSTIYKQGLGVN 593
DB 517 KRLGDVISVQCPVQATVTLRSMR--VPGSETMCSYRPLVSPFSFINDTKTYEGQLGTD 575
```

```
QY 594 NEILLTTTLETCOENTYFYFOAKTDMYIKVYKVEHLKTVPLSSIITLDTFIALNFTLEN 653
DB 576 NEIFLTKWTEVCQATSQYFQSGNHHIYNDVHHFKTIELDGIATLQTFISUNSLIEN 635
QY 654 VDFKVIELYTRDEKRLSNVDFDIETMREYNYAQRVSGLRKDLDD-LSTNRNQFVDAFGS 712
DB 636 IDFASLELYSRDEQRASNVFDLEGIREFYFQAQNIAGLRKOLDNAVSNCRNQFVDGLGE 695
QY 713 LMDDLGAVGQTVVNAVSGVATLFPSSIVTGFINFIPKPPGGMMLMIIVIGVLFAIYFLTKK 772
DB 696 LMDSLGSGQSITNLVSTVGLFSSLSVSGFISPFKPNPFGGMLILVLVAGVILVISLTR 755
QY 773 TKIYETAPIKMIYPEIDKLKEREKSE---IAPISSEELERIVLAMIHCOONSHMETKTR 829
DB 756 TROMSQOPVOMLYPGIDELAQOQHASGEGPGINPISKTELQAIMLA--LHEQNOEQRAAQ 813
QY 830 KOPKDSILTRAQNMRLKR 847
DB 814 RAAGPSVASRALQAARDR 831

RESULT 26
US-10-055-364-46
Sequence 46; Application US/10055364
GENERAL INFORMATION:
APPLICANT: Patience, Clive
TITLE OF INVENTION: Gamma Herpes Virus DNA and Methods of Use
FILE REFERENCE: 61750-379
CURRENT APPLICATION NUMBER: US/10/055,364
CURRENT FILING DATE: 2002-01-23
PRIOR APPLICATION NUMBER: US/09/612,204
PRIOR FILING DATE: 2000-07-07
PRIOR APPLICATION NUMBER: US/60/142,736
PRIOR FILING DATE: 1999-07-08
PRIOR APPLICATION NUMBER: US/60/168,532
PRIOR FILING DATE: 1999-12-02
NUMBER OF SEQ ID NOS: 55
SOFTWARE: Patentin version 3.0
SEQ ID NO 46
LENGTH: 829
TYPE: PRT
ORGANISM: Epstein-Barr virus
US-10-055-364-46

Query Match 40.6%; Score 1844.5; DB 26; Length 829;
Best Local Similarity 46.1%; Pred. No. 4.2e-164;
Matches 367; Conservative 148; Mismatches 240; Indels 41; Gaps 12;

QY 81 STFPYRVCSAGVDVFRFQTDHVCDD-ASDMVHSEGLLIYKQNIIPFMRVKRYKV 139
DB 44 TSPFPRVCELSHGDLFRFSDIOCPFGTRENHTEGLLMVFKDNIIPYSFKVRSYTKIV 103
QY 140 TTSTVYNGIYSDSTINQHTFYKSIPEWETEKMDTIYOCFNSLRNTGNNLLTYVDRDIN 199
DB 104 TNLIYNGWYADSVTNHREKFSVDSYETQMDTIYOCYNAVXKTKDGLTRVYVDRDGVN 163
QY 200 MTVFLQPDVGTDPVKRYGQPELYLEPGFWGYSRRRTTVNCELMDMFARSNPPDFV 259
DB 164 ITVNLKPTGGLANGVRRYASQTELYDAPGMLIWTYRTTIVNCLITDMMAKSNSPDDFFV 223
QY 260 TATGDTVEMSPFWSGEGDDHKNMHEKFPVSVINNYKVVDYQNGRTVPLGKTRIFLDREE 319
DB 224 TTTGQTVMSPFVDGK--NKETTFHERADSFHVRTNYKIVDYDNRGTNPQGERAFLDKGT 281
QY 320 YTLNWK--MSYCPGLTWKAFYNGIOTHEHSGSYHFVANDITASFTTSKEDMKERNNT 378
DB 282 YTLNWK--LENRTAYCPLQHWQTFDSTIATETGKSIHFVDEGTSFVNTTVGIELPDA 339
QY 379 YHCLNEEIKAEIKKAKVNSTHSGYD-LKYFKTDGGLYLVQPLQNRLLDAKN--KL 435
DB 340 FKCIIEQVNTMHEKYEAVQDRTYTKQEAITYFITSGGLLLAWLPLTPRSLATVKNLT 399
```

Qy	436	NNETYS- -RRSRQAESTTDDPMWMTGAGGEYSENSI	473
		: :	
Db	400	TTPTSSPPSPSPAPSAARGSTPAAVLRRRRDAGNATTP- -VPPTAPGKSLGTLNNP	456
		: :	
Qy	474	TVAQOYAYONLIRIRINILEDLSKAWCEHOHAALVMNELSKIAPTSTVMSMIYNRPYSA	533
		: : : : : : : : :	
Db	457	ATVOIQPAYDLSRQRINRMGLDLARAWCLEOKQNMLRELTKINPTTWSSIIYGAVAA	516
		: : : : : : : : :	
Qy	534	KRI GDVISVNCI WV DQT S VSLHKSURLLSASDEKCF SR PPT F K PMND STTYKGQLGVN	593
		: : : : : : : : :	
Db	517	KRLGDVISVSQCVPVNOATVLTRKSMR-VPGSETMCSPLVSFSFINDTKYEGQLGTD	575
		: : : : : : : : :	
Qy	594	NEILLITTYLETQENVEYYFOAKDMYIKYQYEHUKTVPLSSITTLDTFFIALNETLLEN	653
		: : : : : : : : :	
Db	576	NEIFLTKMKTEVCQATSQQYYFQSGNEIHVVNDYHHFKFTIELDGIATLQTFISLNTSLIEN	635
		: : : : : : : : :	
Qy	654	VDFKVIELYTRDKRLSNVFDIETMEREXNYAORVSGLRKDLD-LSTNRNQFYDAFGS	712
		: : : : : : : : :	
Db	636	IDFASIELYSRDSQRASNVPDLGGIFREYNFAQNJAGURKOLDNAVSGNRNQFYVDGLGE	695
		: : : : : : : : :	
Qy	713	LMDDLGAVGQTVNVNAVSGVATLFSSIVTGFIKNPFPGGMLMIIVIGVLPAIYFLTKK	772
		: : : : : : : : :	
Db	696	LMDSLGSVGQSITNLVSTVGGFLSSFVSGFISFFKNPFGGMLILVLVAGVILVLSLTR	755
		: : : : : : : : :	
Qy	773	TKIYETAPIKMVIPEIDKLKEREGKE- -IAPISEEELERIVLAWHIHQOSHMETKTR	829
		: : : : : : : : :	
Db	756	TROMSQOPVOLYPGIDE LAQQAHSAGEGPGINPI SKTELQAI MLA- -LHEQNOEQKRAAQ	813
		: : : : : : : : :	
Qy	830	KDPKDSILTFAQNMLR	845
		: : :	
Db	814	RAAGPSVASRALQAAR	829
		:	
 RESULT 27 US-08-471-913-103			
; Sequence 103, Application US/08471913			
; GENERAL INFORMATION:			
; APPLICANT: Bolognesi, Dani P.			
; APPLICANT: Matthews, Thomas J.			
; APPLICANT: Wild, Carl T.			
; APPLICANT: Barney, Shawn O.			
; APPLICANT: Lambert, Dennis M.			
; APPLICANT: Petteway, Stephen R.			
; APPLICANT: Langlois, Alphonse J.			
; TITLE OF INVENTION: COMPOSITIONS FOR INHIBITION OF MEMBRANE			
; TITLE OF INVENTION: FUSION-ASSOCIATED EVENTS, INCLUDING EPSTEIN-BARR VIRUS			
; NUMBER OF SEQUENCES: 209			
; CORRESPONDENCE ADDRESS:			
; ADDRESSEE: Pennie & Edmonds			
; STREET: 1155 Avenue of the Americas			
; CITY: New York			
; STATE: New York			
; COUNTRY: USA			
; ZIP: 10036-2711			
; COMPUTER READABLE FORM:			
; MEDIUM TYPE: Floppy disk			
; COMPUTER: IBM PC compatible			
; OPERATING SYSTEM: PC-DOS/MS-DOS			
; SOFTWARE: Patent In Release #1.0, Version #1.30			
; CURRENT APPLICATION DATA:			
; APPLICATION NUMBER: US/08/471,913			
; FILING DATE: 07-JUN-1995			
; CLASSIFICATION: 435			
; ATTORNEY/AGENT INFORMATION:			
; NAME: Coruzzi, Laura A.			
; REGISTRATION NUMBER: 30,742			
; REFERENCE/DOCKET NUMBER: 7872-030			
; TELECOMMUNICATION INFORMATION:			
; TELEPHONE: (212) 790-9090			
; TELEFAX: (212) 869-9741/8864			
; TELEX: 66141 PENNIE			
; INFORMATION FOR SEQ ID NO: 103:			


```

399 TTPTSPSPSPAPSAARGSTPAALVLRRLRRDAGNATTP---VPPTAPGKSLGTLNPP 455
474 TVAQOVAYDNLRIINNILEDLSKAWCREOHRALVWNLKINPTSVMSMIYNRPVSA 533
456 ATVQIOFAYDSLRQINRMLGLDLARAWCLEQKQNMVLRKINPTTVNMSIYGRAVAA 515
534 KRIGDVISVNCIIVVDQTSVSLHSLRLSLASDEKCFSRPPVTFKFWNDSTIYKQGLGVN 593
516 KRLGDVISVOCVQVNVQATVLRKSMR-VPGSETWCYSRPLVSFINDTKTYEGQGLTD 574
594 NEILLTTTLETQENTYFYFOAKTDMYIKYKNEHLKTVPLSSITLDTFIALNFTLLEN 653
575 NEIFLTKMTVECVQATSYFYFQSGNEIHVYNDYHFKTIELDGIATLQTFISLNTSLIEN 634
654 VDFKVIETYRDEKLSNVFDEIETMREYNYAQRVSLGRKDLDD--LSTNRQFVDAFGS 712
635 IDFASLELYSRDEQRASNVFDELEGIFREYNFQAQNIAGLRKDLNNAVSGNRNQFVDGLGE 694
713 LMDDLGAVGQTVNVNAVSGVATLFPSSIVTGFINFKNPFGGMLMIIVVIGLVFAIYFLTKK 772
695 LMDSLGSGVQGITNLVSTVGGFLSLSVSGFISFFKNPFGGMLLVLVAGVILVISLRR 754
773 TKIYETAPIKMIYPEIDKLKREKGE---IAPISEEELERIVLAMIHQNSHMETKTR 829
755 TRQMSQOPVQMLYPCIDELAQHASGEGPGINPISKTELQAIMLA--LHEQNOEQKRAAQ 812
830 KDPKDSILTRAQNMURKR 847
813 RAAGPSVASRALQAARDR 830

RESULT 31
US-08-484-223D-103
; Sequence 103, Application US/08484223D
; GENERAL INFORMATION:
; APPLICANT: Barney, Shawn O.
; APPLICANT: Lambert, Dennis M.
; APPLICANT: Petteway, Stephen R.
; TITLE OF INVENTION: ISOLATED PEPTIDES DERIVED FROM HUMAN
; TITLE OF INVENTION: IMMUNODEFICIENCY VIRUS TYPES 1 AND 2 CONTAINING
; TITLE OF INVENTION: FUSION INHIBITORY DOMAINS
; FILE REFERENCE: 7872-029-999
; CURRENT APPLICATION NUMBER: US/08/484,223D
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: 08/470,896
; PRIOR FILING DATE: 1995-06-06
; PRIOR APPLICATION NUMBER: 08/360,107
; PRIOR FILING DATE: 1994-12-20
; PRIOR APPLICATION NUMBER: 08/255,208
; PRIOR FILING DATE: 1994-06-07
; PRIOR APPLICATION NUMBER: 08/073,028
; PRIOR FILING DATE: 1993-06-07
; NUMBER OF SEQ ID NOS: 272
; SOFTWARE: Fast-SEQ for Windows Version 4.0
; SEQ ID NO 103
; LENGTH: 856
; TYPE: PRT
; ORGANISM: Human Immunodeficiency Virus
US-08-484-223D-103

Query Match 40.4%; Score 1839; DB 8; Length 856;
Best Local Similarity 46.18; Pred. No. 1.5e-163;
Matches 368; Conservative 148; Mismatches 240; Indels 42; Gaps 13;

QY 81 STFPYRVCSAGVDVFRFOTDHYCPD-ASDMVHSEGILLIYKONIIPIPFMRVRYKRVV 139
DB 44 TSFPYRVCSAGVDVFRFOTDHYCPD-ASDMVHSEGILLIYKONIIPIPFMRVRYKRVV 103
QY 140 TTSTVYNGIYSDITNOHTYKSIETPWEKMTIYQCFNSLRNLGNGNLLTVYDRDDIN 199
DB 104 TNLIVNGWADSVTRNHEEKFSVDSYETDQMTIYQCYNAVXKTKDGLTRVYVDRDGVN 163
QY 200 MTVFLQPDVGVTDVVKRYGSGQPELYLEPGWFGWSYRRRTTVNCELMDMFARSPPDFV 259

```

```

164 ITVNLKPTGGLANGVRYASQIETLYDAPGWLWITYRTITVNCILITDMAKSNPDPFFV 223
260 TATGDTVEMSPFWSGDEDDHENKMKHFKPWFVSVNNKVVQYQNRGTVPGLKTRIFLDREE 319
224 TTTGQTVEMSPFVDGK--NKETFERADSPHVRTNYKIVDYDNRGTPNQGERAFLDKGT 281
320 YTLISWEKHLKN-MSYCPILTLWKAFYNGIQTEHSGSVHYFVANDITASFTTSKEDMKENNT 378
282 YTLISWK--LENRTAYCPLOHMOTFDSTIATETETGSIHFVTDGTSFVNTTVGIELPDA 339
379 YHCLNEIEIKAEIEKKYAKVNSTHSHKYGD-LKYFKTGGGLYLVWQPLIQNLRLDKN--KL 435
340 FKICIEOVN-KTHEKYEAVQDRTYKQEBALTYFITSGGLLLAWLPLTPRSLATVKNLTTEL 398
436 NNETY-----RRSRQAESTTDPMMEMTNGAGGYSSENSI 473
399 TTPTSPSPSPAPSAARGSTPAALVLRRLRRDAGNATTP---VPPTAPGKSLGTLNPP 455
474 TVAQOVAYDNLRIINNILEDLSKAWCREOHRALVWNLKINPTSVMSMIYNRPVSA 533
456 ATVQIOFAYDSLRQINRMLGLDLARAWCLEQKQNMVLRKINPTTVNMSIYGRAVAA 515
534 KRIGDVISVNCIIVVDQTSVSLHSLRLSLASDEKCFSRPPVTFKFWNDSTIYKQGLGVN 593
516 KRLGDVISVOCVQVNVQATVLRKSMR-VPGSETWCYSRPLVSFINDTKTYEGQGLTD 574
594 NEILLTTTLETQENTYFYFOAKTDMYIKYKNEHLKTVPLSSITLDTFIALNFTLLEN 653
575 NEIFLTKMTVECVQATSYFYFQSGNEIHVYNDYHFKTIELDGIATLQTFISLNTSLIEN 634
654 VDFKVIETYRDEKLSNVFDEIETMREYNYAQRVSLGRKDLDD--LSTNRQFVDAFGS 712
635 IDFASLELYSRDEQRASNVFDELEGIFREYNFQAQNIAGLRKDLNNAVSGNRNQFVDGLGE 694
713 LMDDLGAVGQTVNVNAVSGVATLFPSSIVTGFINFKNPFGGMLMIIVVIGLVFAIYFLTKK 772
695 LMDSLGSGVQGITNLVSTVGGFLSLSVSGFISFFKNPFGGMLLVLVAGVILVISLRR 754
773 TKIYETAPIKMIYPEIDKLKREKGE---IAPISEEELERIVLAMIHQNSHMETKTR 829
755 TRQMSQOPVQMLYPCIDELAQHASGEGPGINPISKTELQAIMLA--LHEQNOEQKRAAQ 812
830 KDPKDSILTRAQNMURKR 847
813 RAAGPSVASRALQAARDR 830

RESULT 32
US-08-485-551-103
; Sequence 103, Application US/08485551
; GENERAL INFORMATION:
; APPLICANT: Bolognesi, Dani P.
; APPLICANT: Matthews, Thomas J.
; APPLICANT: Wild, Carl T.
; APPLICANT: Barney, Shawn O.
; APPLICANT: Lambert, Dennis M.
; APPLICANT: Petteway, Stephen R.
; APPLICANT: Langlois, Alphonse J.
; TITLE OF INVENTION: METHODS FOR INHIBITION OF MEMBRANE
; TITLE OF INVENTION: FUSION-ASSOCIATED EVENTS, INCLUDING INFLUENZA VIRUS
; TITLE OF INVENTION: TRANSMISSION
; NUMBER OF SEQUENCES: 209
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible

```

```

Db      695  LMDSLGVSQSIINLVSTVGLFSSLVSGFISPFKNPFGGMLILVLVAGVVILISLRR 755
Qy      773  TKIYETAPIKMIYPEIDKLKEREKSE---IAPISSEBELERIVLAMHIHQONSHMETKTR 829
Db      755  TROMSQOPVQMLYFGIDELAQOAHSGGPGINPISTKTELQAIMLA--LHEQNEQOKRAAQ 812
Qy      830  KDPKDSILTRAQNWLRKR 847
Db      813  RAAGPSVASRALQAARDR 830

RESULT 33
US-08-487-266-103
; Sequence 103, Application US/08487266
; GENERAL INFORMATION:
; APPLICANT: Bolognesi, Dani P.
; APPLICANT: Matthews, Thomas J.
; APPLICANT: Wild, Carl T.
; APPLICANT: Barney, Shawn O.
; APPLICANT: Lambert, Dennis M.
; APPLICANT: Petteway, Stephen R.
; APPLICANT: Langlois, Alphonse J.
; TITLE OF INVENTION: METHODS FOR INHIBITION OF MEMBRANE
; TITLE OF INVENTION: FUSION-ASSOCIATED EVENTS, INCLUDING HIV TRANSMISSION
; NUMBER OF SEQUENCES: 209
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/487,266
; FILING DATE: 07-JUN-1995
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Coruzzi, Laura A.
; REGISTRATION NUMBER: 30,742
; REFERENCE/DOCKET NUMBER: 7872-025
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-9741/8864
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 103:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 856 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
US-08-487-266-103

Query Match 40.4%; Score 1839; DB 8; Length 856;
Best Local Similarity 46.1%; Pred. NO. 1.5e-163;
Matches 368; Conservative 148; Mismatches 240; Indels 42; Gaps 13;

Qy      81  STFPYRCVCSAGVGDFYRFQTDHVCDD-ASDMVHSSGILLIYKNIITPEMRVRKRYKVV 139
Db      44  TGFPPFRCVLSHGGDUFRRSSDIQCPSFGFRENHTEGLMLVFKDNIILIPYSFKVRSYTKIV 103
Qy      140  TTSTVYNGIVSDITNQHTFYKSIETPEWTEKMDTIYQCFSNLRINTCGNLLTYVDRDDIN 199
Db      104  THILIINGWYADSVTNRHEEKFSVDSYETDMDTIYQCYNVAVKTKDGLTRVYVDRDGVN 163
Qy      200  MTVFLQDPQVTPDVVKRYGSGQPELYLEPGFWGWSYRRRTTVCNCELMDMWFARSNPPFFVF 259
Db      164  ITVNLKPTGLANGVARYASOTELYDPAGWLIWYTRTRTVVNCCLITDMMAKSNPSPPFFVF 223

```

```
QY 260 TATGDTVEMSPFWSGEDDHENKMKHKEKPFVSVINNYKVVDYQNRGTVPGLCKTRIFLDREE 319
DB 224 TTGTQVEMSPFYDGK--NKETHRADSHFVRTNYKIVDYDNRGTNPQGERAFDLKGT 281
QY 320 YTLSEWKLKN-MSYCPLTLWKAIFYNGIQTEHSGSYHFVANDITASFTTSKEDMKBFNTT 378
DB 282 YTLSSK--LENRTAYCPLQHWQTFDSTIATETGKSHFVTDGTSSTPVTNTTVGIELPDA 339
QY 379 YHCLNEBEIKAEIKKYAKUNSTHISKYGD-LKYFKTDGGLYLVWQPLIQNRLLDAKN--KL 435
DB 340 FKCIIEQVN-KTHEKYEAVQDRYTKQEAITFYITSGLLLAWLPLTPRSLATVKNLTTEL 398
QY 436 NNETYSS-----RSRROAESTTDPMMETGNGAGGEYSSENSI 473
DB 399 TTPTSSPPSPSPAPSAARGSTPAVLRRRRRDAGNATTP---VPPTAPGKSLGTINNP 455
QY 474 TVAQVQYADNLRIRINNILEDLSKAWCREQHRAALVWNLKINPTSVMSMITYNRPVSA 533
DB 456 ATVQIQFAYDSLRRQINRMGLDLARAWCLEQKQNMVLRRLTKINPTVMSSYIGKAVAA 515
QY 534 KRIGDVISVNCIVVDQTSVLSLHKSRLLSASDEKCFSPRPVTFKFNWDSITYKGOLGVN 593
DB 516 KRLGDVISVQCQVPVNOATVTLRKSMR-VPGSETMCSYRPLVSFSFINDTKTYEGQLGTD 574
QY 594 NEILLTTTLETQCENTEYFQAKTDMYIYKNYEHKLTVPPLSSITTLDTFIALNFTLLEN 653
DB 575 NEIFLTCKMTEVCQATSYQYFQSGNEIHVNDYHFKTIELDGIATLQTFISLNTSLIEN 634
QY 654 VDFKVIELTRDEKRLSNVFDIETMFREYNYAQRVSGRLKDLID-LSTNRNQFVDAFGS 712
DB 635 IDPASLELYSRDEQASNVFDLEGI FREYNFQAQNIAGRLKDLNNAVSNRQNFVDGLGE 694
QY 713 LMDDLGAQGTVMNAVSGVATLFSSIVTGFINFKNPFQGMMLIIVVIGVLFALYELTKK 772
DB 695 LMDSLGVSQGSITNLAVTSGVGLFSSLSVSGFISFPKPNFGMLIILVAVGVVILVSLTR 754
QY 773 TKIYETAPIKMIYPEIDKLKEREKSE---IAPISEELERIVLAMHIHQONSHMETKTR 829
DB 755 TROMSQOPVOMLYPGIDELAQHRAAGGPGINPLSKTELQAIMLA---LHEQNOEQKRAAQ 812
QY 830 KDPKDSILTAQNMRLRX 847
DB 813 RAAGPSVASRALQAARD 830

RESULT 34
US-08-487-266A-103
; Sequence 103, Application US/08487266A
; GENERAL INFORMATION:
; APPLICANT: Bolognesi, Dani P.
; APPLICANT: Matthews, Thomas J.
; APPLICANT: Wild, Carl T.
; APPLICANT: Barney, Shawn O.
; APPLICANT: Lambert, Dennis M.
; APPLICANT: Petteway, Stephen R.
; APPLICANT: Langlois, Alphonse J.
; TITLE OF INVENTION: METHODS FOR INHIBITION OF MEMBRANE
; TITLE OF INVENTION: FUSION-ASSOCIATED EVENTS, INCLUDING HIV TRANSMISSION
; NUMBER OF SEQUENCES: 239
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
```

```
; APPLICATION NUMBER: US/08/487,266A
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Coruzzi, Laura A.
; REGISTRATION NUMBER: 30,742
; REFERENCE/DOCKET NUMBER: 7872-025
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-9741/8864
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 103:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 856 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; US-08-487-266A-103

Query Match 40.4%; Score 1839; DB 8; Length 856;
Best Local Similarity 46.1%; Pred. No. 1.5e-163;
Matches 368; Conservative 148; Mismatches 240; Indels 42; Gaps 13;

QY 81 STFPYRVCASGVGVFRFQTDHVCDD-ASDMVHSEGIILLIYKQNIIPFMFRVKYRKVV 139
DB 44 TSFPFVVCESLHGDLFPRESSDIQCFSGTRENHTEGLLMVFKDNIIPYSFKVRSYTKIV 103
QY 140 TTSVTVNGIYSDISITNQHTFYKSIPEWTEKMDTIYOCFNSRLNTGGNLLTYVDRDDIN 199
DB 104 TNLIIYNGWYADSVTNRHEEKFSVDSYETQDMDTIYQCVNAVAMTKDGLTRVVVDRDGVN 163
QY 200 MTVFLOPVDGVTDPVKYCSQPELYLERGFWGVSRRRTTUNCELMDMFARSNPDPDFV 259
DB 164 ITVNLKPTGGLANGVRVYASQTELYDAPGLWLTWYTRTRTVNCLITDMMAKSNPDPDFV 223
QY 260 TATGDTVEMSPFWSGEDDHENKMKHKEKPFVSVINNYKVVDYQNRGTVPGLCKTRIFLDREE 319
DB 224 TTGTQVEMSPFYDGK--NKETHRADSHFVRTNYKIVDYDNRGTNPQGERAFDLKGT 281
QY 320 YTLSEWKLKN-MSYCPLTLWKAIFYNGIQTEHSGSYHFVANDITASFTTSKEDMKBFNTT 378
DB 282 YTLSSK--LENRTAYCPLQHWQTFDSTIATETGKSHFVTDGTSSTPVTNTTVGIELPDA 339
QY 379 YHCLNEBEIKAEIKKYAKUNSTHISKYGD-LKYFKTDGGLYLVWQPLIQNRLLDAKN--KL 435
DB 340 FKCIIEQVN-KTHEKYEAVQDRYTKQEAITFYITSGLLLAWLPLTPRSLATVKNLTTEL 398
QY 436 NNETYSS-----RSRROAESTTDPMMETGNGAGGEYSSENSI 473
DB 399 TTPTSSPPSPSPAPSAARGSTPAVLRRRRRDAGNATTP---VPPTAPGKSLGTINNP 455
QY 474 TVAQVQYADNLRIRINNILEDLSKAWCREQHRAALVWNLKINPTSVMSMITYNRPVSA 533
DB 456 ATVQIQFAYDSLRRQINRMGLDLARAWCLEQKQNMVLRRLTKINPTVMSSYIGKAVAA 515
QY 534 KRIGDVISVNCIVVDQTSVLSLHKSRLLSASDEKCFSPRPVTFKFNWDSITYKGOLGVN 593
DB 516 KRLGDVISVQCQVPVNOATVTLRKSMR-VPGSETMCSYRPLVSFSFINDTKTYEGQLGTD 574
QY 594 NEILLTTTLETQCENTEYFQAKTDMYIYKNYEHKLTVPPLSSITTLDTFIALNFTLLEN 653
DB 575 NEIFLTCKMTEVCQATSYQYFQSGNEIHVNDYHFKTIELDGIATLQTFISLNTSLIEN 634
QY 654 VDFKVIELTRDEKRLSNVFDIETMFREYNYAQRVSGRLKDLID-LSTNRNQFVDAFGS 712
DB 635 IDPASLELYSRDEQASNVFDLEGI FREYNFQAQNIAGRLKDLNNAVSNRQNFVDGLGE 694
QY 713 LMDDLGAQGTVMNAVSGVATLFSSIVTGFINFKNPFQGMMLIIVVIGVLFALYELTKK 772
DB 695 LMDSLGVSQGSITNLAVTSGVGLFSSLSVSGFISFPKPNFGMLIILVAVGVVILVSLTR 754
QY 773 TKIYETAPIKMIYPEIDKLKEREKSE---IAPISEELERIVLAMHIHQONSHMETKTR 829
```

```
Db 755 TRMSQPPVQMLYPGIDELAQHASGEGPGINPISKTELQAIMLA--LHEQNEQKRAAQ 812
Qy 830 KDPKDSILTRAQNMRLKR 847
Db 813 RAAGPSVASRALQAARDR 830

RESULT 35
US-08-487-355-103
; Sequence 103, Application US/08487355
; GENERAL INFORMATION:
; APPLICANT: Bolognesi, Dani P.
; APPLICANT: Matthews, Thomas J.
; APPLICANT: Wild, Carl T.
; APPLICANT: Barney, Shawn O.
; APPLICANT: Lambert, Dennis M.
; APPLICANT: Petteway, Stephen R.
; APPLICANT: Langlois, Alphonse J.
; TITLE OF INVENTION: METHODS FOR INHIBITION OF MEMBRANE
; TITLE OF INVENTION: FUSION-ASSOCIATED EVENTS, INCLUDING HEPATITIS B VIRUS
; TITLE OF INVENTION: TRANSMISSION
; NUMBER OF SEQUENCES: 209
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/487,355
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Coruzzi, Laura A.
; REGISTRATION NUMBER: 30,742
; REFERENCE/DOCKET NUMBER: 7872-027
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-9741/8864
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 103:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 856 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
US-08-487-355-103

Query Match 40.4%; Score 1839; DB 8; Length 856;
Best Local Similarity 46.1%; Pred. No. 1.5e-163;
Matches 368; Conservative 148; Mismatches 240; Indels 42; Gaps 13;

Qy 81 STFPYRVCSAGSGVGVFRFQTDHVCPCD-ASDMVHSEGIILYKONIIPMFVRVRYKRV 139
Db 44 TSFPRVELSHGDLFRSSDIQPCSGTRENHTGELLMVFKONIIPYFKVRSYTKIV 103
Qy 140 TTSTVYNGIYSITNOHTFFYKSIPEWTERKMDTIYQCNSLRNLNTGGNLLTYVDRDDIN 199
Db 104 TWILIYNGWYADSVNTRHEEKFSVDSYETDQMDTIYQCNVAVMTKOGLTRVYVDRDGVN 163
Qy 200 MTFVLPQVDGVTVPDKRVQSGQELYLEPCGFWSGTRRTTNCLEMDMFARNSPPDFPV 259
Db 164 ITWNLKPTGGLANGVRRYASOTELYDAPGWLITVTRTTRTNCLITDMMAKSNSPDPFV 223
Qy 260 TATGDTVEMSPWSEDDHENKMKPWFVSVNNYKVVQYQNRCTVPLGKTRIFLDREE 319

Db 224 TTTGQTVEMSPFYDGK--NKETFHERADSPHVRNYKIVDYDNRGTNPQGERRAFLDKGT 281
Qy 320 YTLSEWEXHLKN-MSYCPCLTLWKAFYNGIQTEHSSSYHFPVANDITASFTTSKEDKKEFWNTT 378
Db 282 YTLSSWK--LENRTAYCPLQHWQTFDSTIATETGSKIHFVTDDEGTSSFTVNTTVGIELPDA 339
Qy 379 YHCLNEEIKAEIEKKYAKVNSTHSHKYGD-LKYFXTDGGLYLVWQPLIQNRLLDARN--KL 435
Db 340 FKCIIEQVN-KTHEKEYAVQDRYTKQEAITYFITSGLLLAWLPLPRSLATVKNUTEL 398
Qy 436 NNETY-----RRSRQAESTTDPMMEMTGNAGGEYSSNSI 473
Db 399 TPTTSSPPSPSPSAARGSTPAALVLRRRRRDAGNATP---VPPTAPGKSLGTLNPN 455
Qy 474 TVAQVOYAYDNLRIINNILEDLSKANCROHRAALVWNLKSKINPTSVSMIYNRPVSA 533
Db 456 ATVOIQFAYDSLRRQINRMGLDLARAWCLEQKQNMVLRRLTKINPTTVMSSIYGKAVAA 515
Qy 534 KRIGDVISVNCIVVDOTSVSLHKSRLLSASDEKCFSRPPVTFKFMNDSTIYKQGLGVN 593
Db 516 KRLGDVISVOCVPVQATVTLKSMR-VPGSETWCYSRPLVSFSFINDTKTYEGQLGTD 574
Qy 594 NEILLTTTLETQENTYFYFQAKTDMYIYKYNHKLTVPLSSITTLDTFIALNFTLLEN 653
Db 575 NEIFLTKMTVEVCQATSQYFYQSGNEIHVYNDYHHFKTIELDGIATLQTFISLNTSLIEN 634
Qy 654 VDFKVIELYTRDEXKLSNVFDIETMFREYNYAQRVSGRLRKDLDD-LSTNRNQFVDARGS 712
Db 635 IDFASLELYSDEQRASNVFDEGIFREYNFQAQNIAGLRKLDNAVSNRGNQFVDGLGE 694
Qy 713 LMDDLGAVGQTVNVNAVSVATLFPSSIVTGFNFTKNPFGGMILIVIGLVFAIYFLTKK 772
Db 695 LMDSLGVSQGSITNLVSTVGGLFSSLVSGFISFPNPGMILIVLVAGVVILVISLRR 754
Qy 773 TKIYETAPIKMIYPEIDKLKEREKSE---IAPISEEELERIVLAMHIHQONSHMETKTR 829
Db 755 TRMSQPPVQMLYPGIDELAQHASGEGPGINPISKTELQAIMLA--LHEQNEQKRAAQ 812
Qy 830 KDPKDSILTRAQNMRLKR 847
Db 813 RAAGPSVASRALQAARDR 830

RESULT 36
US-08-487-355A-103
; Sequence 103, Application US/08487355A
; GENERAL INFORMATION:
; APPLICANT: Bolognesi, Dani P.
; APPLICANT: Matthews, Thomas J.
; APPLICANT: Wild, Carl T.
; APPLICANT: Barney, Shawn O.
; APPLICANT: Lambert, Dennis M.
; APPLICANT: Petteway, Stephen R.
; APPLICANT: Langlois, Alphonse J.
; TITLE OF INVENTION: METHODS FOR INHIBITION OF MEMBRANE
; TITLE OF INVENTION: FUSION-ASSOCIATED EVENTS, INCLUDING HEPATITIS
; TITLE OF INVENTION: B VIRUS TRANSMISSION
; NUMBER OF SEQUENCES: 273
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/487,355A
```

```

Db 755 TRMSQOPVOMLYPGIDELAQOQHASGEGPINSKTELQAIMLA--LHEQNOQEKRAAQ 811
Qy 830 KDPKDSILTRAQNMLRKR 847
: : : : :
Db 813 RAAGPSVASRALQAARDR 830

RESULT 37
US-08-919-600-103
: Sequence 103, Application US/08919600
: GENERAL INFORMATION:
: APPLICANT: Boignesi, Dani P.
: APPLICANT: Matthews, Thomas J.
: APPLICANT: Wild, Carl T.
: APPLICANT: Barney, Shawn O.
: APPLICANT: Lambert, Dennis M.
: APPLICANT: Petteway, Stephen R.
: APPLICANT: Langlois, Alphonse J.
: TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR INHIBITION
: OF INVENTION: OF MEMBRANE FUSION-ASSOCIATED EVENTS, INCLUDING HIV
: TITLE OF INVENTION: TRANSMISSION
: NUMBER OF SEQUENCES: 273
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Pennie & Edmonds
: STREET: 1155 Avenue of the Americas
: CITY: New York
: STATE: New York
: COUNTRY: USA
: ZIP: 10036-2711
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/919,600
: FILING DATE:
: CLASSIFICATION:
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/470,896
: FILING DATE: 06-JUN-1995
: ATTORNEY/AGENT INFORMATION:
: NAME: Coruzzi, Laura A.
: REGISTRATION NUMBER: 30,742
: REFERENCE/DOCKET NUMBER: 7872-020
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (212) 790-9090
: TELEFAX: (212) 869-9741/8864
: TELEX: 66141 PENNIE
: INFORMATION FOR SEQ ID NO: 103:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 856 amino acids
: TYPE: amino acid
: STRANDEDNESS:
: TOPOLOGY: unknown
: MOLECULE TYPE: protein
: US-08-919-600-103

Query Match 40.4%; Score 1839; DB 13; Length 856;
Best Local Similarity 46.1%; Pred. No. 1.5e-163;
Matches 368; Conservative 148; Mismatches 240; Indels 42; Gaps 13;

Qy 81 STPYRVCSAGVDYRFOTDHWCPD-ASDMVHSEGIILLYKONIIPFMRVKYKVV 139
: : : : :
Db 44 TSPFPRVCELSHGGDUFRRSSDTCQPSFGREHNHTEGLLMVFKDNIIPYGFKVRSYTKIV 103
: : : : :
Qy 140 TTSTVNGIYSDSITNQHTYKYEIPEWETEKMDTIYOCFNSRLNTGNLLTYVDRDDIN 199
: : : : :
Db 104 TNLINGWVADSVTWRHEKFSVDSYETQMDTIYQCYNAVAKMKDGLTRVYVDRGVN 163
: : : : :
Qy 200 MTVFLQPDVGVTVDVKRYGSOPELYLEPGWFWGVSRRRTTVNCCLMDMFAARSNPFFDFV 259
: : : : :
Db 164 ITVNLKPTGGLANGVRYASOTELXPAGWLIWYRTRITVNCCLITDMMAKNSPFFDFV 223
: : : : :

```

```
QY 260 TATGDTVEMSPFMSGDDHDKMHEKPFVSVNNYKVVDYQNRGTVPGLKTRIFLDREE 319
Db 224 TTTGQTVEMSPFYDCK--NKETPHERADSFHVRTNYKIVDYNRGTPNPOGERRAFLDKGT 281
QY 320 YTLSEKHLKX-MSYCPDLTKAFYNGIOTHSYGHFVANDITASTFTTSKEDMKFNTT 378
Db 282 YTLSEKHLKX-MSYCPDLTKAFYNGIOTHSYGHFVANDITASTFTTSKEDMKFNTT 378
QY 379 YHCLNEEIKAEIEKKYAKVNSTHSHKYGD-LKYFKTDGGLVLMQPLQNLRLDKN--KL 435
Db 340 FKCIIEQVN-KTHEKEAVQDRYTKGQEAITYFITSGGLLLAWLPTRSLATVKNLTTEL 398
QY 436 NNETYS-----RRSRQAESTTDPMMEMTNGAGGEYSSENSI 473
Db 399 TPTTSSPPSPSPAPSAARGSTPAVLRRLRRRRDAGNATP---VPPTAPGKSLGTNNP 455
QY 474 TVAQOVAYDNLRIIRINILEDLSKAWCREQHRALVWNLSEKINPTSVMSMIYRNPVSA 533
Db 456 ATVOIQFAYDSLRRQINRMGLDLARAWCLEQRQNNVRLTKINPTTVMSIYKAVAA 515
QY 534 KRIGDIVSNCIIVDDOTSVSLHKLRLLSASDEKCFSPRPVTEKFMNDSTIYKGLGVN 593
Db 516 KRLGDIVSQCVPVNOATVTLRKSMR-VPGSETWCYSRPLVSVSFINDTKTYEGOLGTD 574
QY 594 NEILLTTTLETCEQNTTEYFQAKTDMYIKYVEHLKTVPLSSITTLDTFIALNFTLLEN 653
Db 575 NEIFLTKMTEVCQATSOYFYFQSGNEIHVYNDYHHFKTIELDGIATLQTFISLNTSLIEN 634
QY 654 VDFKVIELTRDEKRLSNVFDIETMPREYNYAQRVSGRLKDLDD-LSTNRNQVDAFGS 712
Db 635 IDFASLELYSRDEQRASNVFDEGIFREYNFQAQNIAGLRKDLDNAVNSGRNQVVDGLGE 694
QY 713 LMDDLGAVGQTVNVAVSGVATLFSSIVTGFINFKNPFPGGMLIIVIGVLFAYFLTKK 772
Db 695 LMDSLGSGVQSTINLVSTVGGVLFSSVSGFISFFKNPFPGGMLIIVLAGVILVLSLRR 754
QY 773 TKIYETAPIKMITYPEIDKLRKREGKSE---IAPISEELERIVLAWHIHQOQSHMETKTR 829
Db 755 TFQMSQOPVQMLYPGIDELAQOHAQSGEGPGINPISKTELQALMLA--LHEQNOEQKRAQ 812
QY 830 KDPKDSILTRAQNLKR 847
Db 813 RAAGPSVASRALQAARD 830
```

RESULT 38

```
US-09-502-445-103
; Sequence 103, Application US/09502445
; GENERAL INFORMATION:
; APPLICANT: Bolognesi, Dani P.
; APPLICANT: Matthews, Thomas J.
; APPLICANT: Wild, Carl T.
; APPLICANT: Barney, Shawn O.
; APPLICANT: Lambert, Dennis M.
; APPLICANT: Petteway, Stephen R.
; APPLICANT: Langlois, Alphonse J.
; TITLE OF INVENTION: METHODS FOR INHIBITION OF
; TITLE OF INVENTION: MEMBRANE FUSION-ASSOCIATED EVENTS, INCLUDING HIV
; NUMBER OF SEQUENCES: 244
; CORRESPONDENCE ADDRESS:
; ADDRESS: Pennie & Edmonds LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
```

```
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/502,445
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/487,266
; FILING DATE: 07-JUN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Coruzzi, Laura A.
; REGISTRATION NUMBER: 30,742
; REFERENCE/DOCKET NUMBER: 7872-025
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-9741/8864
; TELETYPE: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 103:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 856 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; US-09-502-445-103
```

```
Query Match 40.4%; Score 1839; DB 19; Length 856;
Best Local Similarity 46.1%; Pred. No. 1.5e-163;
Matches 368; Conservative 148; Mismatches 240; Indels 42; Gaps 13;
QY 81 STFPYVCSAGVGDFRFDHVCDD-ASDMVHSEGLLIYKQNIIPFMFRVKRYKV 139
Db 44 TSFPFRVCESLSSGDLFRFSSDIQCPSPGTRENHTEGLLMVFKDNIIPYSFKVRSYTKIV 103
QY 140 TTSVYNGIYSDSITNOHTFYKSIPEWTEKMDIYQCFNSRLNTGNTLLTYVDRDDIN 199
Db 104 TNLIIYNGWYADSVNRHEEKFSVDSETDQMDIYQCYNAVAKMTKGLTRVYVDRDGVN 163
QY 200 MTVFLQPDVGTDPVKRYGQPELYLEPGFWGVSRRRTVNCCLMDMFARSNPDPDFV 259
Db 164 ITVNLKPTGGLANGVRRYASQTELYDAPGWLITWYRTTIVNCLITDMAKSNPDPDFV 223
QY 260 TATGDTVEMSPFMSGDDHDKMHEKPFVSVNNYKVVDYQNRGTVPGLKTRIFLDREE 319
Db 224 TTTGQTVEMSPFYDCK--NKETPHERADSFHVRTNYKIVDYNRGTPNPOGERRAFLDKGT 281
QY 320 YTLSEKHLKX-MSYCPDLTKAFYNGIOTHSYGHFVANDITASTFTTSKEDMKFNTT 378
Db 282 YTLSEKHLKX-MSYCPDLTKAFYNGIOTHSYGHFVANDITASTFTTSKEDMKFNTT 378
QY 379 YHCLNEEIKAEIEKKYAKVNSTHSHKYGD-LKYFKTDGGLVLMQPLQNLRLDKN--KL 435
Db 340 FKCIIEQVN-KTHEKEAVQDRYTKGQEAITYFITSGGLLLAWLPTRSLATVKNLTTEL 398
QY 436 NNETYS-----RRSRQAESTTDPMMEMTNGAGGEYSSENSI 473
Db 399 TPTTSSPPSPSPAPSAARGSTPAVLRRLRRRRDAGNATP---VPPTAPGKSLGTNNP 455
QY 474 TVAQOVAYDNLRIIRINILEDLSKAWCREQHRALVWNLSEKINPTSVMSMIYRNPVSA 533
Db 456 ATVOIQFAYDSLRRQINRMGLDLARAWCLEQRQNNVRLTKINPTTVMSIYKAVAA 515
QY 534 KRIGDIVSNCIIVDDOTSVSLHKLRLLSASDEKCFSPRPVTEKFMNDSTIYKGLGVN 593
Db 516 KRLGDIVSQCVPVNOATVTLRKSMR-VPGSETWCYSRPLVSVSFINDTKTYEGOLGTD 574
QY 594 NEILLTTTLETCEQNTTEYFQAKTDMYIKYVEHLKTVPLSSITTLDTFIALNFTLLEN 653
Db 575 NEIFLTKMTEVCQATSOYFYFQSGNEIHVYNDYHHFKTIELDGIATLQTFISLNTSLIEN 634
QY 654 VDFKVIELTRDEKRLSNVFDIETMPREYNYAQRVSGRLKDLDD-LSTNRNQVDAFGS 712
Db 635 IDFASLELYSRDEQRASNVFDEGIFREYNFQAQNIAGLRKDLDNAVNSGRNQVVDGLGE 694
QY 713 LMDDLGAVGQTVNVAVSGVATLFSSIVTGFINFKNPFPGGMLIIVIGVLFAYFLTKK 772
```

Db 695 LMDSLGSGQITNLVSTVGLFSSLVSGFISPFKNPFGMLILVLVAGVILVSLTR 754
Qy 773 TKIVETAPIKMIYPEIDKLKEREKSE---IAPISEEELERIVLAMIHOONSHMETKTR 829
Db 755 TROMSQOQVOMLYPGIDELAQQHASGEGPGINPISKTELQAIMLA--LHEQNOEQKRAAQ 812
Qy 830 KDPKDSILTRAQNMLRKR 847
Db 813 RAAGPSVASRALQAARDR 830

RESULT 39
US-10-267-682-103
; Sequence 103, Application US/10267682
; GENERAL INFORMATION:
; APPLICANT: Bolognesi, Dani P.
; Mathews, Thomas J.
; Wild, Carl T.
; Barney, Shawn O.
; Lambert, Dennis M.
; Petteway, Stephen R.
; Langlois, Alphonse J.
; TITLE OF INVENTION: COMPOSITIONS FOR INHIBITION OF
; MEMBRANE FUSION-ASSOCIATED EVENTS, INCLUDING HIV
; TRANSMISSION
; NUMBER OF SEQUENCES: 239
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/267,682
; FILING DATE: 08-Oct-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/484,223A
; FILING DATE: 07-JUN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Coruzzi, Laura A.
; REGISTRATION NUMBER: 30,742
; REFERENCE/DOCKET NUMBER: 7872-029
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-9741/8864
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 103:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 856 amino acids
; TYPE: amino acid
; STRANDEDNESS: <Unknown>
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 103:
US-10-267-682-103

Query Match 40.4%; Score 1839; DB 28; Length 856;
Best Local Similarity 46.1%; Pred. No. 1.5e-163;
Matches 368; Conservative 148; Mismatches 240; Indels 42; Gaps 13;

Qy 81 STPPYRVCSAGVGDVFRFQTDHVCDD-ASDMVHSEGLIILYKONIIIPFMRVRKRVKVV 139
Db 44 TSPFPRVCSHGDLPFRSSDIQCPFGTRENHTEGLLMVFKDNIIPYGFKVRSTKIV 103
Qy 340 TTSTVYNGIVSDSITNQHTFYKSETPWETEKMDTIYQCFNSRLNTGGLNTYVDRDDIN 199

Db 104 TNLINYNWADSVTRNEHEKFSVDSYETQMDTIYQCYNAVMTKDGRLTRVYVDRGVN 163
Qy 200 MTVFLQPDVGTDPVKRYGSOPELYLEPGFWGYSRRRTTVNCCLMDPMFARSNPPFFV 259
Db 164 ITVNLKPTGGLANGVRRYASOTELYDAPGWLITWTRTRTTVNCCLITDMAKSNPPFFV 223
Qy 260 TATGDTVEMSPFWSGEDDHNKMKHPWFSVNNYKVVDYQNRGTIVPLGKTRIFLDREE 319
Db 224 TTTGQTVEMSPFYDGK--NKETPHERADS FHVRTNYKIVDYDNRGTPOGERRAFLDKGT 281
Qy 320 YTLISWEKHLKN-MSYCPCLTLWKAFYNGIQTEHSGSYHFVANDITASTTISKEDKKEFNIT 378
Db 282 YTLISWK--LENRTAYCPLQHWOTFDSTIATETGKSIHFVDEGTSSFTVTTVIGIELPDA 339
Qy 379 YHCLNEIEKAEIEKKYAKVNSTHSGYGD-LKYFKTDGGLYLWQPLIQNLRLDKN--KL 435
Db 340 FKCTEQVN-KTHEKVEAVQDRYTKGQEAITYFTTSGLLALLMLPLTPRSLATVKNLTTEL 398
Qy 436 NNETYS-----RRSRQAESTTDPMMEMTGNGAGGEYSSENSI 473
Db 399 TPTSPSPSPSPSPAPSAARGSTPAAVLRRRRRDAGNATTP---VPPTAPGKSLGTLNPP 455
Qy 474 TVAQVOVAYDNLRIRINNILEDLSKAWCRQHRRAALVMNELSKINPTSVMSMIYNRPVSA 533
Db 456 ATVQIQFAYDSLRRQINRMGLDLARAWCLEQKQNMVLRELTKINPTTVMSSYICKAVAA 515
Qy 534 KRIGDVISVNCIIVDQTSVSLHKSURLLSASDEKCFSPRPVTFKFNMDSTIYKQGLGVN 593
Db 516 KRLGDIVISVQCPVNAQATVTLRKSMT-VPGSETMCSRPLVSFSFINDTKTYEGQLGTD 574
Qy 594 NEILLTTTLETQENTYFYFOAKTDMYIYKNEHKLTVPLSSITTLDTIALNFTLEN 653
Db 575 NEIFLTKMKTEVCQATSQYFQSGNEIHVYNDYHFKFTELDGIATLQTTISLNTSLIEN 634
Qy 654 VDFKVELYTRDEKRLSNVDFDIETMPREYNVYQRYSGRLKRLDLD-LSTNRNORFVDAFGS 712
Db 635 IDFALELYSRDEQRASNVFDLEGI FREYNFQANIAGRLKDLNDVNSGRNORFVDCGLGE 694
Qy 713 LMDDLGAVGQTVVNAVSGVATLFSSTIVTGFINFKNPFQGMGLMIIVVIGLVFAIYFUTKK 772
Db 695 LMDSLGSGQITNLVSTVGLFSSLVSGFISPFKNPFGMLILVLVAGVILVSLTR 754
Qy 773 TKIVETAPIKMIYPEIDKLKEREKSE---IAPISEEELERIVLAMIHOONSHMETKTR 829
Db 755 TROMSQOQVOMLYPGIDELAQQHASGEGPGINPISKTELQAIMLA--LHEQNOEQKRAAQ 812
Qy 830 KDPKDSILTRAQNMLRKR 847
Db 813 RAAGPSVASRALQAARDR 830

RESULT 40
US-10-267-748-103
; Sequence 103, Application US/10267748
; GENERAL INFORMATION:
; APPLICANT: Bolognesi, Dani P.
; Mathews, Thomas J.
; Wild, Carl T.
; Barney, Shawn O.
; Lambert, Dennis M.
; Petteway, Stephen R.
; Langlois, Alphonse J.
; TITLE OF INVENTION: COMPOSITIONS FOR INHIBITION OF
; MEMBRANE FUSION-ASSOCIATED EVENTS, INCLUDING HIV
; TRANSMISSION
; NUMBER OF SEQUENCES: 239
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA

Db 668 YSLAGIKDLDNTIDYNRDLRLVQDLSDMMADLGDIGRSVVNVSSVTVFFSSIVTGFIKF 727
Qy 746 IKNPFGGMLMIIVIGVFAIYFLTKTKIYETAPIKMIYPIEIDKLKREG----KSBIA 801
Db 728 FTNPLGGIFILLIIGIIFLVVLRNRSQFHDAPIKMLYPSVENYAAQAPPPYSASPP 787
Qy 802 PISEELERIVLAMH-IHQONSHMETK-TRKDPKDSILTRAQNMRLR-KRSGYSNL 853
Db 788 AIDKEEIKRILLGMHVHQBEEAKQKLTNSGP--TLWQKATGFLNRNRKGYSQL 840

RESULT 43
US-09-338-326-17
; Sequence 17, Application US/09338326
; GENERAL INFORMATION:
; APPLICANT: Rose, Timothy M.
; APPLICANT: Bosch, Marnix L.
; APPLICANT: Strand, Kurt
; TITLE OF INVENTION: GLYCOPROTEIN B OF THE RFHV/KSHV
; OPERATING SYSTEM: SUBFAMILY OF HERPES VIRUSES
; NUMBER OF SEQUENCES: 100
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morrison & Foerster
; STREET: 755 Page Mill Road
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304-1018
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/338,326
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/720,229
; FILING DATE: 26-SEP-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Schiff, J. Michael
; REGISTRATION NUMBER: 40,253
; REFERENCE/DOCKET NUMBER: 29938-20002.00
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 813-5600
; TELEFAX: (415) 494-0792
; TELEX: 706141
; INFORMATION FOR SEQ ID NO: 17:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 849 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-338-326-17

Query Match 39.4%; Score 1792; DB 17; Length 849;
Best Local Similarity 42.4%; Pred. No. 4.1e-159;
Matches 354; Conservative 158; Mismatches 271; Indels 52; Gaps 13;

Qy 34 SPPNTATWSTPLTGHVGYTHDSHGERGNENRDESEQNKIYGPSTFPYRVCSAGSV 93
Db 43 TPAQAPTEPTPLS-----TWNRCFEY-----FRVCGVAAT 75
Qy 94 GDVFRFQTDHVCPSADSMVHSEGIILYKONIIPFMFRVRYKRVKVVTTSTVYNGIYSDSI 153
Db 76 GETFRFLDKTSPSTODKKHVEGILLVYKINIVPIFKIRYKRIITQITWRGLTSSV 135
Qy 154 TNOHTFYKSIPEWETEKMDTYQCNSRLNTGNNLLTYVDRDDINMTVFLQPDGVTPD 213
Db 136 TCKPEMATQAEWEVGDPSIYQYCNYSATVMVNVNRYVYVDRDGVNKTNIRPVDGLTGN 195

Qy 214 VKRYGSOPELYLEPGWFGWSYRRRTTNCLELMDMFARSNPDPFVVTATGDTVMSPFWS 273
Db 196 IQRYSQPTLLYSEPGWMPGFVVRTTVNCEIVDMVARSDPYNATATAGDSLELSPQT 255
Qy 274 GEDHENKMHK-PWFVSVINNYKVQYQNGRTVPLGTRIFLDRREEYTLSEWKLKMS 332
Db 256 FDNTSQCTAPKRAMRVREKNYKFDVNNRGTPAGOSRTFLETSPSATYSMTATROTA 315
Qy 333 YCPILTWKAFYNGTQTEHSGSYHFVANDITASFTTSKEDMKEPNTYHCLNEEIKAEK 392
Db 316 TCDLVHMKTFPRAIQTAHEHSYHFVANEVATFNTPLTEVENFTSTYSCVSDQINKTISE 375
Qy 393 KYAKVNSTHSGYDLKYFKTDGGLYLVWQPLIQNLRLDAKNKLANE-----TYSRRSRQ 447
Db 376 YIQKLNNYSVASGKTQYFKTDGGLYLVWQPLHEPEIEDIDEDSDPEPTPAFPKSTRKRE 435
Qy 448 AESTTDPMMETGNGAGGEYSSENSITVAQVQYAYDNLRIRINILLEDLSKAWCREQURA 507
Db 436 AADNGNSTSEVS-----KGSNPLITAQIQFAYDKLTTSVNNVLEELSRACREQVRD 488
Qy 508 ALVNEELSKINPTSMVMYINRYPVSAKRGIDVISVNCIVVDOTSVSLHLSLRLSASDE 567
Db 489 TLMYELSKVNPTSVNGAIYKPVAAVYGVDAISVTDCIYVDQSSVNIHQSLR-LQHDKT 547
Qy 568 KCFSRPPVTFKFMNDSTIYKQGLGVNNEILLTTLTYLETQCENTEYVYFOAKTDMYIKNYE 627
Db 548 TCYSRPRVTFKFINSTDPLTGQLGPRKEIILSNINIECTKDESEHYFVGEVYIYKNYI 607
Qy 628 HLKTVPLSSITLDTFFIALNFTLLENDFKVIELYTRDEKRL-SNVFDIETMFREYNYA 686
Db 608 FEELNLSSITLDTFFIALNISFIENIDFKVELYSSTERKLASSVFDIESMFREYNYT 667
Qy 687 QRVSLRKDLDD-LSTNRNQVDAFGSLMDLGAQGVQTVNVAVSGVATLSSIVTGFINF 745
Db 668 YSLAGIKDLDNTIDYNRDLRLVQDLSDMMADLGDIGRSVVNVSSVTVFFSSIVTGFIKF 727
Qy 746 IKNPFGGMLMIIVIGVFAIYFLTKTKIYETAPIKMIYPIEIDKLKREG----KSBIA 801
Db 728 FTNPLGGIFILLIIGIIFLVVLRNRSQFHDAPIKMLYPSVENYAAQAPPPYSASPP 787
Qy 802 PISEELERIVLAMH-IHQONSHMETK-TRKDPKDSILTRAQNMRLR-KRSGYSNL 853
Db 788 AIDKEEIKRILLGMHVHQBEEAKQKLTNSGP--TLWQKATGFLNRNRKGYSQL 840

RESULT 44.
US-08-804-439-15
; Sequence 15, Application US/08804439
; GENERAL INFORMATION:
; APPLICANT: Rose, Timothy M.
; APPLICANT: Bosch, Marnix L.
; APPLICANT: Strand, Kurt
; TITLE OF INVENTION: GLYCOPROTEIN B OF THE RFHV/KSHV
; OPERATING SYSTEM: SUBFAMILY OF HERPES VIRUSES
; NUMBER OF SEQUENCES: 113
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morrison & Foerster
; STREET: 755 Page Mill Road
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304-1018
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/804,439
; FILING DATE:
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Schiff, J. Michael

```

; REGISTRATION NUMBER: 40,253
; REFERENCE/DOCKET NUMBER: 29938-20002.20
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 813-5600
; TELEFAX: (415) 494-0792
; TELEX: 706141
; INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 874 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-804-439-15

Query Match      39.3%; Score 1787; DB 12; Length 874;
Best Local Similarity 43.5%; Pred. No. 1.3e-158;
Matches 367; Conservative 159; Mismatches 248; Indels 70; Gaps 20;

QY   34 SPNTATWSTESPLTGHYTHDSSHGRCNNENRDEEQNKNIYGSPSTFFPYRVCSA-SG 92
DB   58 SPQNT---STSKEPSTDNQGTSTPT-----IPTVDDTASKNFY-----KYRVCSASSS 102

QY   93 VGDVFRFOTDHVCPPASDMVHSEGIILLYKONIIPFMFRVKRYKVVTTSVTVNGIYSDS 152
DB   103 SGELEFRFDLDQCPTDKKKHVEGILLVLKNIVPIFKVRKYRKIATSVTVYRGWSQA 162

QY   153 ITNQHFTFYKSIEPWETEKMDTIYQCFNSLRNLTCGNLLTYVDRODINMTVFLOPVGVTP 212
DB   163 VTNRDDISRAIPYNELSMIDRTYHCFSAMATVINGILTNYIDROSENKSVLPQVAGLTE 222

QY   213 DVKRYGSQPELYLEGWFGWGRYRRRTTNCELMDFARNSPPFPFFVTATGDTVMESPFW 272
DB   223 NINRYFSQPLIIYAEPGWPGIYRVRTVNCVVDMYARSVEPYTHFITALGDTIEISPF 282

QY   273 SGEDDENKM-----HEKPMFVSVINNVYDVONRGTVPLGKTIRFLDRREYTL 322
DB   283 -----HNNSQCTTGNSTSRDATKVM---IEENHOTVDYERRGH-PTDKRIFLKDEEYTI 333

QY   323 SWEKHLKNMSYCPLTLWKAFYNGIOETHSGSVHFVANDITASFSTSKEDMKEF--NP-TY 379
DB   334 SWAEADRERAI CDVFIWKTFPRAIQTIHNESPHFVANEVTASFSTNOEETELRGTEIL 393

QY   380 HCLNEIKAEIKKAKVNASTHSKYGLKFYTGGGLYLVMOQLTONRLDLDAKN---KLN 436
DB   394 NCWNSTINETLEETVKFKFKSHIRGEVKYYKTNGSLFLIQAMKPLNLSHTNYTIERN 453

QY   437 NETYSRRSRQEAESTTDPMEMWTGNGAGEYSSSENSITVAOVAYDNMLRIRINNILEDL 496
DB   454 NKTKGNSRKRSYDVT-----KTFQAGK-----LSTAQVOYAYDHLRTSMNHILEEL 500

QY   497 SKAWCREOHRALAVNELSKINPTSVMISINYRPVSARKIGDIVSVSNICIVVDQTSVSLH 556
DB   501 TKTWCREQKNDLMMYELSKINPVSMAIYGPVAVKMGDAFMVSECINVDQASVNIH 560

QY   557 KSLRLLSASDEK-CFSRPPVTETKFNWDSTIYKGQGVANNEILLTTYLETCOEINTYVFO 615
DB   561 KSNR---TDDPKVCYSRPLVTETKFNVNSTATFRGQLGTNEILLTNHWETECRPTADHYFF 617

QY   616 AKTDMYIYKNYEHLKTVPLSSITTLDLTIALNFTLLENVDFKVIELYTRDEKRILSNVFI 675
DB   618 VKNMTHYKDYKVFVKMTDNTNISTLDLTFLTLNLTFTIDNIDEKVELYSETERKMASALDL 677

QY   676 ETMFREYNYAQBVSGRLKDL---IDLSTNRNQFVDAFGSLMDDLGAQGCVTVNAVSGVA 732
DB   678 ETMFREYNYTOKLASRLDELONTIDL--NDRDLVKDISSEMADLGDTGKVVVNTFSGIV 735

QY   733 TLFSSIVTGFINKPNPFCGMIMIIWIGVLFAIYFLTKTKIYETAPIKMIYPEIDKLK 792
DB   736 TVFGSIVGFGVFVFTNPVGVIILLIIVVPVPIVSRPTNNWNEAPIKMIYINIDKAS 795

QY   793 EREGKSEIAPISEEEELERVILAMWHIQQNQSHMETKTRKD PKD---SILTRAQNMLKRSG 849

```


THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 28, 2003, 15:30:15 ; Search time 31 Seconds
(without alignments)
1276.284 Million cell updates/sec

Title: US-10-055-364-24
Perfect score: 4547
Sequence: 1 MAGSLKLRGSLVLAFLYQV.....KRGYSNLKNAESVEMLNTL 865

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 242630 seqs, 45739658 residues

Total number of hits satisfying chosen parameters: 242630

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Pending Patents_AA_New.*

- 1: /cgn2_6/ptodata/2/paa/PCT_NEW_COMB.pep.*
- 2: /cgn2_6/ptodata/2/paa/US06_NEW_COMB.pep.*
- 3: /cgn2_6/ptodata/2/paa/US07_NEW_COMB.pep.*
- 4: /cgn2_6/ptodata/2/paa/US08_NEW_COMB.pep.*
- 5: /cgn2_6/ptodata/2/paa/US09_NEW_COMB.pep.*
- 6: /cgn2_6/ptodata/2/paa/US10_NEW_COMB.pep.*
- 7: /cgn2_6/ptodata/2/paa/US60_NEW_COMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1813	39.9	854	5	US-09-350-841A-1589
2	844	18.6	904	1	PCT-US03-11231-18
3	836.5	18.4	904	1	PCT-US03-30301-56
4	789.5	17.4	943	6	US-10-670-695-14
5	244.5	5.4	227	1	PCT-US03-30301-54
6	148.5	3.3	1431	6	US-10-679-063-18001
7	135.5	3.0	1638	1	PCT-US03-26635-1
8	130.5	2.9	774	6	US-10-425-114A-43519
9	130.5	2.9	1276	6	US-10-425-114A-8
10	128	2.8	1001	1	PCT-US03-20460-10
11	128	2.8	1193	6	US-10-679-063-18223
12	126.5	2.8	3899	1	PCT-US02-18638A-4
13	126.5	2.8	3907	1	PCT-US02-18638A-2
14	126.5	2.8	3917	1	PCT-US02-18638A-8
15	126.5	2.8	3925	1	PCT-US02-18638A-6
16	126	2.8	912	6	US-10-679-063-18203
17	124.5	2.7	353	6	US-10-425-114A-55825
18	124	2.7	250	6	US-10-670-695-16
19	124	2.7	1038	1	PCT-US03-27401-460
20	124	2.7	1038	6	US-10-472-928-4532
21	123	2.7	526	6	US-10-472-928-3638
22	122.5	2.7	1286	6	US-10-205-516A-22
23	122	2.7	1664	1	PCT-US02-24459-102
24	122	2.7	1738	1	PCT-US02-24459-100
25	121.5	2.7	836	6	US-10-679-063-11928
26	121	2.7	1147	6	US-10-679-063-18214

27	121	2.7	2004	1	PCT-US03-27401-364	Sequence 364, App
28	121	2.7	2004	6	US-10-472-928-2306	Sequence 2306, App
29	120.5	2.7	846	6	US-10-679-063-18007	Sequence 18007, A
30	120	2.6	793	6	US-10-679-063-18197	Sequence 18197, A
31	120	2.6	1732	6	US-10-679-366-2	Sequence 2, Appli
32	118.5	2.6	1280	1	PCT-US03-27401-274	Sequence 274, App
33	118.5	2.6	1280	6	US-10-472-928-364	Sequence 364, App
34	117.5	2.6	928	1	PCT-US03-30720-1209	Sequence 1209, App
35	117.5	2.6	1030	6	US-10-425-114A-62748	Sequence 62748, A
36	117	2.6	880	6	US-10-330-773-956	Sequence 956, App
37	117	2.6	889	6	US-10-679-063-18166	Sequence 18166, A
38	116	2.6	1301	6	US-10-205-516A-18	Sequence 18, Appl
39	116	2.6	1881	1	PCT-US03-27401-316	Sequence 316, App
40	116	2.6	1881	6	US-10-472-928-1236	Sequence 1236, App
41	116	2.6	2497	6	US-10-679-063-18067	Sequence 18067, A
42	114.5	2.5	1169	6	US-10-472-928-4712	Sequence 4712, App
43	113.5	2.5	1431	6	US-10-679-063-20989	Sequence 20989, A
44	113.5	2.5	1518	1	PCT-US03-06962-40	Sequence 40, Appl
45	113.5	2.5	1617	6	US-10-679-063-18152	Sequence 18152, A

ALIGNMENTS

RESULT 1

US-09-350-841A-1589
; Sequence 1589, Application US/09350841A
; GENERAL INFORMATION:

- ; APPLICANT: Jeffs, Peter;
 - ; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR INHIBITION OF MEMBRANE
 - ; FILE REFERENCE: 7872-066-999
 - ; CURRENT APPLICATION NUMBER: US/09/350.841A
 - ; CURRENT FILING DATE: 1999-07-09
 - ; NUMBER OF SEQ ID NOS: 1946
 - ; SOFTWARE: PatentIn Ver. 2.1
 - ; SEQ ID NO 1589
 - ; LENGTH: 854
 - ; TYPE: PRT
 - ; ORGANISM: Human immunodeficiency virus type 1
- US-09-350-841A-1589

Query Match	39.9%	Score 1813;	DB 5;	Length 854;
Best Local Similarity	46.0%	Pred. No. 8.4e-91;		
Matches	367;	Conservative 147;	Mismatches 240;	Indels 44; Gaps 15;
Qy	81	STFPYRVCASGVGDVFRFQTDHVCPD-ASDMVHSEGLLIYKONIIPFMRFRKRVKW	139	
Db	44	TSPFPRVCELSSHGDLFRFSSDIQCPSGFTRENHTEGLLMVFKDNIIPYSEKRSYTKIV	103	
Qy	140	TTSTVYNGIYSDSITNOHTFYKSIPEWETEKMDTIYQCFNSLRNLNTGNNLLTYVDRDIN	199	
Db	104	TNLIYNGWADSVTNRHEEKFSVDETQMDTIYQCYNAVKMTKDLTRVYVDRDGVN	163	
Qy	200	MTVFLQPDVGVTPDVKIKYGSOPELYLPFGWFGSYRRRTTVCNCELMDFARSNPPDFPV	259	
Db	164	ITVNLKPTGGLANGVRVYASOTELYDAPGLIWTYRTTVCNCLITDMAKSNPPDFPV	223	
Qy	260	TATGDTVMGPFWSGEDDHENKMKHEKPFWSVSNVYKVDYQNRGTVPGLKTRIFLDREE	319	
Db	224	TTTGQTVEMSPFYDGK--NKETFERADSFHVRTNYKIVDYNRGTNPQGERAFDKGT	281	
Qy	320	YTLSEWELKN--MSYCPDLTLWAFYNGIOTHSYSHFVANDITASTTTSKEDKENTT	378	
Db	282	YTLSEWELKN--MSYCPDLTLWAFYNGIOTHSYSHFVANDITASTTTSKEDKENTT	339	
Qy	379	YHCLNEBIKAEIKYAKVNSTHSGYGD-LKYFKTDGGLVLMQPLIQNLIDAKN--KL	435	
Db	340	FKCEEQVN--KTHEKYEAVQDRYTKGQEAITYFITSGLLLAWLPLTPRSUATVKNLT	398	
Qy	436	NNETYS-----RRSRRAQESITDPMEMTNGAGGEGYSSNSI	473	
Db	399	TTPTSSPSPSPAPSAARGSTPAVLRRRRRDAGNATTP---VPPTAPGKSLGLNNP	455	


```
QY 81 STFPYRVCSAGVDVFRFQTDHVCPCDASD-MVHSEGIILYKQNIIPFMFRVRYKRV 139
Db 115 -----VCPPTGATVVOEQPRCPTRPEQGNQYTGIAVFKENIAPYKFKATWYKDV 168
QY 140 TTSTV-----YNGIYSDSTNQHTFYKSTIEPWE--TEKMDTIYQCFNSURLNTGNL 189
Db 169 TVSQWFMGHRYSQFMGIFED-----RAPVFEVIDKINAKGVCRCSTAKYVRNNDL 219
QY 190 LTVYDRDDINMTVFLQPDGV-----TPDKRYGSOPELYLEPGFWGYSRRRTTVN 241
Db 220 TTAHRRDDHETDMLKPAWAATRTSRGWHITDLEKNPSRVEAF-----HRYGTITV 270
QY 242 CELMDFAHSNPPDFPFVATGDTVEMSPFWG--GEDDH--ENKQHEKPMFVSVINNY-KV 298
Db 271 CIVEEDARSVPYDFEVLATGDFVYMSFFYREGSHTEHTSYAADRKFQVDGYFARDL 330
QY 299 DYONRGTVPLGTRIFLDREBYTLSEKHLKMSVCLTLWKAIFYNGIQTEHSGSYHFA 358
Db 331 TTAKARATAP--TTRNLLTTPKFTVAMDVWPKRPSVCTMTKQOEVDMLRESEYGGSF 388
QY 359 NDITASFTT-----SKEDMKEFTTYHCLNEEIKAEIEKKYA-KVNSTHSKYGDLK 410
Db 389 DAISTFTTNLTPELPSRVDLGD-----CIKDARDAMDRIFARRYNATHIKVQPOQY 442
QY 411 KTDGGLYLWQPLIQNLRLDAKNKLNNETYSRRSROAESTTDDPMEMTNGAGGEYSSE 470
Db 443 LANGGFLIAYQPLLSNTLAE-----LYVREHLREQSRKPNPTPPPPGASANASVERIK 498
QY 471 NSITVAQVQYADNRIIRINNILEDLSKAWCREQRAALVNLSEKINPTSVMSMYNRP 530
Db 499 SSIEFARLOFTYHNIQRHVNMDLGRVAIAWCELOHNEHLTLNNEARKLPNALIASATV 558
QY 531 VSAKRIGDVISVNCIVVDQTSVLSHLKLSLASDEKCFSPPTVTFKFMNDSTIYKGL 590
Db 559 VSARMGLDGMVASTCPVAADNVIVONSMT--ISSRPGACYSLPSFRYEDQGPLVEGOL 617
QY 591 GYNNELITTTTLETQCENTEYVFOAKTDMYIKNVYEHKTVPLSSITTLDTFIALNFTL 650
Db 618 GENNELRLTRDAIEPCTVGHRRYFTFGGYVYFESAYSHQSLRADITTVTFIDLNTM 677
QY 651 LENVDKUIELYTRDEKLSNVDFIETMFREYNYAQRVSGRLKDLDLSTNRNQF--VD 708
Db 678 LEDHEFVPLEVYTRHEIKDSGLLDYTEVQRNQLHDLRFADIT-DTVIHADANAAMFAGLG 736
QY 709 AFGSLMDLGL-AVGOTVNAVSGVATLPSIIVTGFINIKNPFQGLMLIIVIGVLFAY 767
Db 737 AFFEGMGLGRVAGVKVWGIIVGGV-----SAVSGVSVFMSNPFGLAVGLVLVLAGLAAAF 792
QY 768 FLTKTKIYETAPIKMIYPEIDK-LK-----EREGKSEIAPISE-EELERIVLA 814
Db 793 FAFRYVMRLQSNPKALYPLTTKELKNPTNPDSAGEGEGDFDEAKLAAREMIRYNAL 852
QY 815 MHIHQONSHMETKTRDKPKDSILT-RAQNML---KRRSGYNLKNAE 857
Db 853 VSAMERTEH---KAKKKGTSALLSAKVTDVMVRKRRTNYTQVPNKD 896
```

RESULT 4

```
US-10-670-695-14
; Sequence 14, Application US/10670695
; GENERAL INFORMATION:
; APPLICANT: Jensen, Wayne A.
; APPLICANT: Lappin, Michael R.
; APPLICANT: Rosen, David K.
; APPLICANT: Andrews, Janet S.
; TITLE OF INVENTION: USE OF RECOMBINANT ANTIGENS TO DETERMINE THE IMMUNE
; STATUS OF AN ANIMAL
; FILE REFERENCE: DI-9-1
; CURRENT APPLICATION NUMBER: US/10/670,695
; CURRENT FILING DATE: 2003-09-25
; PRIOR APPLICATION NUMBER: 09/521,738
; PRIOR FILING DATE: 2000-03-09
```

```
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 14
; LENGTH: 943
; TYPE: PRT
; ORGANISM: Feline herpesvirus 1
US-10-670-695-14
```

```
Query Match 17.4%; Score 789.5; DB 6; Length 943;
Best Local Similarity 27.3%; Pred. No. 1.5e-35;
Matches 231; Conservative 151; Mismatches 364; Indels 101; Gaps 27;
```

```
QY 80 PSTFPYRVCSAGVDVFRFQTDHVCPCDAS--DMVHSEGIILYKQNIIPFMFRVRYKRV 138
Db 121 PSTP--YMCPPPSGSTVRLPEPRACPDYKLGKNTFEGIAVIFKENIAPYKFKANIYKN 178
QY 139 VTTSTVNGIYSDSTNQHTFYKSTIEPWE--TEKMDTIYQCFNSURLNTGNLLIYVDRDD 197
Db 179 IIMTVMSGSSYAVTTNRYTDRVPKVEITDLIDRGMCLSKADYVRNNYQFTAFDRDE 238
QY 198 INMTVFLQPDGVTPDKRYGSOPELYLEPGFWG-----YRRRTTVNCLMDM 247
Db 239 DPRELPLKPSSTLSRVR-----GWHTNETYTKIVLDFHSGTSGVNCIVEEV 286
QY 248 PARSNPPDFPFVATGDTVEMSPFWGSD-----DHENKMHKPMFVSVINNYKVVDYQNR 303
Db 287 DARSVVPVDSFAISTGDVVIHMSPPFGLRDGAHVEHTSYSSDR---FQIEGYYPIDLDTD 343
QY 304 GT-VPLGKTRIFLDREBYTLSEKHLKMSVCPITLWKAIFYNGIQTEHSGSYHFAVNDIT 362
Db 344 YTGAPV--SRNPLETPHTVAVNMTPKSGRVCTLAKEWREI--DEMPLMNIGSYRFTAKTIS 400
QY 363 ASFTTSKEDMKEFTTY--HCLNEEIKAEIEKKY-KVNSTHSKYGDLKYPKTDGGLYL 419
Db 401 ATFTSNTSQF-EINRIKLGDCATKEAABADRIYKSYKSTHITGTLETYLARGGFLIA 459
QY 420 WQPLIQNL-----LDANKLN--NETYSRRSR-----RQESTTDP 454
Db 460 FRPMISNELAKYINELARSNRTVVDLSALLNPGEIVQTRRSVPNSQHRSSRTIEG 519
QY 455 MMEWTNGAGGEYSSENSITVAQVYADNRIIRINNILEDLSKAWCREQRAALVNL 514
Db 520 GIETVNNAS--LLKTTSSVEFAMLQFAYDYTQAHVNEMLSRITATAWCTLQNRHVLWMTET 577
QY 515 SKINPTSVMSMYNRPVSAKRIGDVISVNCIVVDQTSVLSHLKLSLASDEKCFSPRP 574
Db 578 LKLNPGGVVSNALERRVSARLLGDGAVAVTQCNVSISSGHVYIIONSMR-VTGSSTTCYSRPL 636
QY 575 VTFKFMNDSTIYKQGLGVNNEILLTTTLETQCENTEYVFOAKTDMYIKNVYEHKTVPL 634
Db 637 VSFRLNDSEYIEQOLGENNELVERKLI.EPCTVNNKRYKFGADYVYFEDYAVRVKVP 696
QY 635 SSITTLDTFIALNTLLENVDKVIELYTRDEKLSNVDFIETMFREYNYAQRVSGRLK 694
Db 697 SEIELISAYV-IKSTLLEDREF-LHSSYTRAELDGTGPFDYSEIQRNRQLHALKFYDI-- 752
QY 695 DLLDLSTNRNQFVDAFGSLMDLGLAVGQTVNAVSGVATLPSIIVTGFINIKNPFQGL 754
Db 753 DSVIRVDNNLVIMRGMANFFQGLGDVGAGFGVKVGLGAASAVISTVSGVSSPLNPFGLA 812
QY 755 MIIWVIGVLFAYELTKTKIYETAPIKMIYPEIDKLEREGKSE-----IAP 803
Db 813 VGLLILAGIYVAAFLAYRYISRLRANPMKALYPTVTRNLKQTAKSPASTAGSDSPGVDDF 872
QY 804 SEELER-----IVLAMHIHQONSHMETKTRDKPKDSILT-RAQNMLRKRSK--YS 851
Db 873 DEELKMQAREMIKYMVSLVSAM---EQOEHKAMKKNKGP--AILTSLTNNALRRRGPKYQ 927
QY 852 NLKNAES 858
Db 928 RLNNLDS 934
```

RESULT 5

PCT-US03-30301-54
 ; Sequence 54, Application PC/TUS0330301
 ; GENERAL INFORMATION:
 ; APPLICANT: SYKES, KATHRYN F.
 ; APPLICANT: HALE, KATHERINE S.
 ; APPLICANT: JOHNSTON, STEPHEN A.
 ; TITLE OF INVENTION: METHODS FOR VACCINE IDENTIFICATION AND COMPOSITIONS FOR
 ; TITLE OF INVENTION: VACCINATION COMPRISING NUCLEIC ACID AND/OR POLYPEPTIDE
 ; TITLE OF INVENTION: SEQUENCES OF THE HERPESVIRUS FAMILY
 ; FILE REFERENCE: MCR0:002W0
 ; CURRENT APPLICATION NUMBER: PCT/US03/30301
 ; CURRENT FILING DATE: 2003-09-23
 ; PRIOR APPLICATION NUMBER: 60/412,956
 ; PRIOR FILING DATE: 2002-09-23
 ; NUMBER OF SEQ ID NOS: 116
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID-NO 54
 ; LENGTH: 227
 ; TYPE: PRT
 ; ORGANISM: Herpes Virus
 PCT-US03-30301-54

Query Match 5.4%; Score 244.5; DB 1; Length 227;
 Best Local Similarity 28.3%; Pred. No. 1.1e-06;
 Matches 67; Conservative 52; Mismatches 95; Indels 23; Gaps 8;

Qy 254 PFDFVTATGTVMSDFWS-GEEDH-ENKMKHPFVSVINNY-KVVDYQNRGTVPGLGK 310
 Db 3 PYDFEVLATGDFVMSPFYGYREGSHTHTSYAADRFQVDGFYARDLTTKARATAP--T 60
 Qy 311 TRIFLDREEVTLSEKHLKMSYCPLTLWKAFYNGIOTHSYVHFVANDITASFTT--- 367
 Db 61 TRNLATTPKFTVADWVPKPSVCTMTKQOEVDMLRSEYGGSRFSSDAISTFTTNLT 120
 Qy 368 ----SKEDMKFNTYHCLNEEIKAEIEKKYA-KVNSTHSGYDGLKYFKTDGGLYLWQP 422
 Db 121 EYPLSRVDLGD-----CIGKARDAMDRIFARYNATHIKVQPOYYLANGGFLIAYQP 174
 Qy 423 LIONRLDANKLNNTYRSRRQAESTTDPMMETGNGAGGEYSSENITVAQVQ 479
 Db 175 LLSNTLAE---LYVREHLREQSRKPPNTPPPPGASANASVERIKTSSIEFARLQ 227

RESULT 6

US-10-679-063-18001
 ; Sequence 18001, Application US/10679063
 ; GENERAL INFORMATION:
 ; APPLICANT: Edgerton, Michael D
 ; TITLE OF INVENTION: TRANSGENIC PLANTS WITH IMPROVED PHENOTYPES
 ; FILE REFERENCE: 38-15(52054)B
 ; CURRENT APPLICATION NUMBER: US/10/679,063
 ; CURRENT FILING DATE: 2003-10-02
 ; PRIOR APPLICATION NUMBER: 60/415,758
 ; PRIOR FILING DATE: 2002-10-02
 ; NUMBER OF SEQ ID NOS: 27373
 ; SEQ ID NO 18001
 ; LENGTH: 1431
 ; TYPE: PRT
 ; ORGANISM: Plasmodium yoelii yoelii
 US-10-679-063-18001

Query Match 3.3%; Score 148.5; DB 6; Length 1431;
 Best Local Similarity 18.2%; Pred. No. 0.86;
 Matches 192; Conservative 154; Mismatches 338; Indels 373; Gaps 55;

Qy 50 HYGTHDSHGERGNENR-----DSEQNKNIYVGSPTF 83
 Db 242 NYDNNTYNNVDRNQQNSYVADENETKDNNEEDDDTGTVDNDEEDNDDNDYS 301
 Qy 84 PYRVCSASGVDGVRFF-----QTDHVCPDASDMVHSEGIL-----LIY-KONI 125
 Db 302 QXNQCEVESDTNQIRSNKRYNESIKHI-NESNLKINKESLLKRETYNKRDNIFYIKDOI 360

Qy 126 IPE-----MFRVRKYRKVVTTSTVYNGIYSDSITN-OHTFYKSIETPWT-----EK 170
 Db 361 IPYKKEHNNNIFSLYDSSK---NNNEHNNNYETKFVNYKHGDEKKDEQMLTNVKKERNEE 417
 Qy 171 MDTIYQCFNSL-----RLNTGGNLLTYVDRDDINMTVFLQPDVG-----TPDV 214
 Db 418 FLKKYWEALAKTYSSKDLQFNRLGDEKNILHH-----DINVDKMFRLDKLERSNTODI 473
 Qy 215 KRYGSOPELYLEPGWPGWSYRRRTT-----VNCELMDMFARSPPPDFVTATGDTVEM 268
 Db 474 RNKTNK-----YNNFRHDTTSLSKIRYKDEINDYSEYANKCLE-----GNIYD- 516
 Qy 269 SPWMSGEDD-----HENKMEKHPFVSVINNYKVVDYQNRGTVPGLGTRI--FL 315
 Db 517 -----EDDSYLKRKELGHKELKVIDNP-----ILN-----INHDDK-----KKTNKSFL 557
 Qy 316 DREETLSWEKHLKN-----MSYCPLTLWKAFYNGIOTHSYVHFV 357
 Db 558 DKIKYRKNNELFLKNEHEKYADIQETNKKDKIMKEADILHTKCNFNP---HRGSSIA 614
 Qy 358 A-----NDITASFTT-----SKEDMKF-----NTVH-CLNEIK- 387
 Db 615 KSLSPNNKKNANFNNEAFNLILEKKDKNSDYKQYNIKCFPYDHENTSPHSVRBETKY 674
 Qy 388 -----AEIEKKYAKVNSTHSGYDGLKYFKTDGGLYLWQPLIQNRLD 430
 Db 675 YENKESRKLVDYCDENDIENSIRYNENNIDYNSINETIN-----LREDIKYNEFLN 730
 Qy 431 AKNKLNNETYSRR---SRRQAESTTDPMMETGNGAGGEYSSENS-----ITVAQVQYAY 482
 Db 731 KPDKINSEKFNNSFNDSKQSLKNDSDNKIRDTNKLYYHNNSDNALRSNVIKKESTEKIY 790
 Qy 483 DNLIRINNILEDLSKA-WCREOHRALVWNE--LSKIN--PTSVMYMIYNRPVSAKRI 536
 Db 791 DILESNNMRKPFINKMDYTKYEQNLKNDEDIYTKENKFPDS-RQAFYKNSNSPMKV 849
 Qy 537 GVVISVS-----NCIVV-----DOTSVLSLHSLRLLLSASDEKCFSPRPVTFKF 579
 Db 850 KOETKDSQVEDYNRIRIKYKLMKQEKODEDETDINIDKLSRKMIENNK-----LKE 901
 Qy 580 MNDSTTYKQGLGVNNEILLTITYLET-----COBTEYV-----FOAKTDMVIYKN 625
 Db 902 QNES-----NDELIVTPFYKSKIDKVLKNSEIFERSARATFKQFDVKN-----KN 947
 Qy 626 YHLKTVPLSSITTLDTFTALNPTLLENVDKVEL-----YTRDEKRLSNVFIETWFE 681
 Db 948 FLH-----FSEIESLTKLCYNLE-LPPVDKKILSIYVYKDYDSSKNCNMVMDFROMYWD 1001
 Qy 682 Y-----NYVAQVSGRLKDLDLSTNRNQVDVAFGSLMDDLGAQGTVVNAVSGVATLF 735
 Db 1002 LKQIKKKYPTKNFKIKRNCI-ISRKLQGYD-----Y 1034
 Qy 736 SSIVTGFINFIK-----NPFGGMLMIIVIGVLFAIYFLTKTKTK----- 774
 Db 1035 SSII-NYLSFKILGCGAFGEVHLVEDNICKLYKVVKILKKKKMKMKIKNVNEINVLILYLD 1093
 Qy 775 -----IYETAPIKMIYPEI-----DKLKEREKSEIAPISSEELERIVLAWHHI 818
 Db 1094 HPNIKIFDVYESVNCYIIVMELCEGGLMSKIKPEIFSEIY-IKNIMFOILCAIAYMH 1152
 Qy 819 QQN-SHWETKTRKDPKDSILTRAQNMRLKRSYGNLK 854
 Db 1153 SNIAHKDLK-----PENILFKTDGYDTLK 1177

RESULT 7

PCT-US03-26635-1
 ; Sequence 1, Application PC/TUS0326635
 ; GENERAL INFORMATION:
 ; APPLICANT: INCYTE CORPORATION; BAUGHN, Mariah R.;
 ; APPLICANT: RICHARDSON, Thomas W.; WARQUIS, Joseph P.;
 ; APPLICANT: SWARNAKAR, Anita; Tang, Y. Tom;

```

; APPLICANT: BECHA, Shanya D.; EMERLING, Brooke M.;
; APPLICANT: JIN, Pei; WILSON, Amy D.;
; APPLICANT: YUE, Henry; GIETZEN, Kimberly J.;
; APPLICANT: CHANG, Hsin-Ru; YANG, Yonghong G.;
; APPLICANT: LEE, Soo Yeun; KHARE, Reena;
; APPLICANT: ELLIOTT, Vicki S.; HAPALIA, April J.A.;
; APPLICANT: CHAMLA, Narinder K.; RAMKUMAR, Jayalaxmi;
; APPLICANT: GURURAJAN, Rajagopal; TRIBOULEY, Catherine M.;
; APPLICANT: CHIEN, David; TRAN, Uyen K.;
; APPLICANT: MURAGE, Jaji
; TITLE OF INVENTION: KINASES AND PHOSPHATASES
; FILE REFERENCE: PF-1566 PCT
; CURRENT APPLICATION NUMBER: PCT/US03/26635
; PRIOR FILING DATE: 2003-08-25
; PRIOR APPLICATION NUMBER: US 60/406,172
; PRIOR FILING DATE: 2002-08-26
; PRIOR APPLICATION NUMBER: US 60/413,910
; PRIOR FILING DATE: 2002-09-25
; PRIOR APPLICATION NUMBER: US 60/414,296
; PRIOR FILING DATE: 2002-09-27
; PRIOR APPLICATION NUMBER: US 60/417,821
; PRIOR FILING DATE: 2002-10-11
; NUMBER OF SEQ ID NOS: 118
; SOFTWARE: PERL Program
; SEQ ID NO 1
; LENGTH: 1638
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No: 7504529CD1
PCT-US03-26635-1

Query Match
Best Local Similarity 3.0%; Score 135.5; DB 1; Length 1638;
Matches 124; Conservative 95; Mismatches 201; Indels 247; Gaps 31;

QY 151 DSITNQHTFKYSIEPWEKMDTIYQCENS-URLMTGGNLLTYVDRDDINMTVFLQPDVG 209
Db 190 DSVHQHLYHVRDIKP-----DNILMDXNGHRLADFSCLXLMDEGTVQSSVAVGTDPY 243
QY 210 VTPDV-----KRYGQPE-----LYLEPGWF-----WG----SYRRRTTV 240
Db 244 ISPEILQAMEDGKGRYPGECDWMSLGVCMYELGETPFYABSLVETYGKIMNHRERQF 303
QY 241 NCELMDMFARSNPPDFVTATGDTVMS-----PWSCEDDHENKMHKPFVSV- 291
Db 304 PAQVTDVSENAKDPIRRLICGHEHRLGQSGIEDFKKHPFSGIDWDNIRNCEAPYIPEVS 363
QY 292 ----INNYKVD--YONRQTVP-----LQGT 311
Db 364 SPDTNSFVDVDDCLKNSTMPPTHTAPSGHHLFPVGTYTSSCVLSDRSCLRVTAQPT 423
QY 312 RI-----FLDREYTLSEWKLKNMSYCPLLTKWAFYNGIQTSHSGSYHFVANDITA 363
Db 424 SLDDLVNQVORTLNNLATEAYERRIKRLEQKLESRKLQESTQTVQALQYSTVDGPLTA 483
QY 364 SFTTSKEDKENTYTHCLNEI--KAELEKKYAKVNSTHSHKYGLD-----KYFTD 413
Db 484 SKDLEIKNLKE---VIEKLRKVQTESHLEQOOLEEAVRQELDAFRQIKAYEKQIKT- 539
QY 414 GGLYLWQPLIQNLLDAKNKLNNTYS-----RRSRROAESTTDPMMEMTNGAGG 465
Db 540 -----LQER--EDLNKLEVITEALAAASKDKRLQRESE----- 572
QY 466 EYSS--ENSIT-VAQVQVAYDNLIRINNILEDLSKAWCREQHRAALVWNLSKINPTSV 522
Db 573 HYSKOLENELEGKQKQISY-----SPGVCSIEHQ-----QETTKLKTDL 613
QY 523 MSMIYNPVSAKRIGDVTISVNCIVVDQTSVSLH-----KSLRLLSASDEKCFSPVPTFK 578
Db 614 KKSIFYEBELSKREG-----IHANEIKNLK-----K 639

US-10-425-114A-43519
; Sequence 43519, Application US/10425114A
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114A
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 43519
; LENGTH: 774
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: 700468168_FLI.pep
US-10-425-114A-43519

Query Match
Best Local Similarity 2.9%; Score 130.5; DB 6; Length 774;
Matches 101; Conservative 78; Mismatches 203; Indels 129; Gaps 20;

QY 434 KLANNE-TYSRRSRQAEASTTDPMMEMTNGCAGGEY--SSENSITVAQVQVAYDNLIRIN 490
Db 30 KTNELAINFNLSSANTTINSLODALSQARSNIFILAAEKNEAEKAYETEINALNAELT 89
QY 491 NILEDLSKAWCREQHRAALVWNLSKIN----PTSVNMSI---YNRPVSAKRIGDVTISVS 543
Db 90 KCLBELDKTHGLQSHSTHYGYLEKLGFMFMDDSLLSLAEYEGKTFNSLR-----D 142
QY 544 NCIVVDQTSVSLHLSKRLLSASDEKCFSPRPVTFKMNDSITYKGQ----- 590
Db 143 MCLIVK----SMHBQLSVKS-----FONDSIVEDSELSSLLSLPDYESFV 183
QY 591 -----GVNNEILLTITYLETCQENTYVFOAKTDWYIYKNYE---HLKTVPLSS 636
Db 184 RELVKAINRKGNIDDTSSFSSTIVEQLSNQTEYLSFLKOLSTYMSNIIIVLRLSLQVS 243
QY 637 ITTLDITFIALNFTLLENVDFKVIELYTRD-----EKRLSNVFDIETMFREYNYAQRV 689
Db 244 NTFART-----LEBHDMLKVELGNKDAHNAQAESEVLSQLDKRAMSSKCIYCVQOI 295
QY 690 SGLRKDLLDI-----STNRNQFVDAFGSLMDDLGAVGQTVVNAV 728
Db 296 EIVFDDMWDVGYAIDLATGSSGISELVTVDLKNEDTSDYSKVAOTLATIDTLKS 355
QY 729 SGVATLFSSIVTGFNF---IKNPFGG-----MLMIIVWIGVLFAIYFLTKTKTIYE 777
Db 356 EKLSITIKGLVITSLDDDFKMLKQAEAAEAETASHEHQLSV-----ERVCMLEKELIKL 408
QY 778 TAPTKMTYPIDKLKEREGSEIAPISEELERI-VLAMHIHQONSHMETKTRKDPKDSI 836
Db 409 DECNMEL-NIQEYKEREG-----ALKARELELLSVHEHTQISADRLGTDNAISKQDEAL 462
```


QY 747 KNPFGLMIIIVIGVLFAYF-----LTKTKIYETAPIKMIYPEIDKLEREG 796
Db 875 KTEPTVKIDNQIKNLPFSQFENNYPDYGFVIKTSKNLESSK-----PEAAKVAAPKS 929
QY 797 KSEIAPISEEBLERIVLANHIHQ 819
Db 930 AAKPVAAPKEQOE-----IHO 945
RESULT 11
US-10-063-18223
; Sequence 18223, Application US/10679063
; GENERAL INFORMATION:
; APPLICANT: Edgerton, Michael D
; TITLE OF INVENTION: TRANSGENIC PLANTS WITH IMPROVED PHENOTYPES
; FILE REFERENCE: 38-15(52054)B
; CURRENT APPLICATION NUMBER: US/10/679,063
; CURRENT FILING DATE: 2003-10-02
; PRIOR APPLICATION NUMBER: 60/415,758
; PRIOR FILING DATE: 2002-10-02
; NUMBER OF SEQ ID NOS: 27373
; SEQ ID NO 18223
; LENGTH: 1193
; TYPE: PRT
; ORGANISM: Plasmodium falciparum 3D7
US-10-063-18223

Query Match 2.8%; Score 128; DB 6; Length 1193;
Best Local Similarity 19.7%; Pred. No. 9.3; Mismatches 263; Indels 214; Gaps 40;
Matches 150; Conservative 135
QY 100 QTDHVCPOASDMVHSEGI--LLIYQNIIPFMRVKRVKVTTSVYVNGIYSDSITNOH 157
Db 475 QNTNDCTYSNDINDYKINNSLIPNEN-----EKQNNKSTTKSLYNNIETHSVNTQNR 528
QY 158 TPKYKIEPWETEKMD--TIYQCFNSLRNTGNGLLTYVDRDDINMTVFLQPVGVTPDVK 215
Db 529 S-----EPHEIDMDMDKIDTKCKDPLRNI-----NERNIDVF--IEALLSELK 570
QY 216 RYSGOPELYLEPGWFGSYR--RTVNCELMDMFARSNPFPD-----FFVTATGDTVM 268
Db 571 LVSSKSDL-----SYNSLCRTYESLKLEDDKKAQESADILKRYFIHDKNNLNE 621
QY 269 SPFWGEDEDDHENKHE-----KPMFVSVINNVKV--VDYQNRGTVPGLKTRIFLDRREVT 322
Db 622 SYLEYKEEN--CKIREIILSIFKYLDDIINNEDISVLVSNNELTPTQDNNMNMNN 679
QY 323 SWEKHLKNNSCYCP--LTLWKAFY-----NGIQTEHSGSYHFVANDITASFTTSK 369
Db 680 KQNIKINKNKKNILCVVKWYLLNIQRIEMVINAL-----SNVTFV--DI-----NK 726
QY 370 EDMKEFNTYHCLNEBEIKA-----EIEKKYAKVNS-----THSKYGDLYFKTD 413
Db 727 NVLEISSIIYIYIELIKKLEDNYSYAIIFHDIINKYCSDDNLEITHDKYEV----- 780
QY 414 GGLYLWQPL-----IQNRLLDAKNNKNNETYSRRSR----- 445
Db 781 -GLNLC-KPLFSSLMMFTNVILLNSNTDVKIKADITRKLFYTNINISRFPMUKISGI 838
QY 446 -----ROAE-----STDPMMEMTNGAGGEYSSENSITVAQVOYAY-----DNL 485
Db 839 LIRILTNKPIEQAKFYFSTFEVIK-----KASDVVKPLIPQLQTCIIKSLNNENL 890
QY 486 RIRINILEDLSKAWCREQRAALVWNL-----SKNPTSM-----SMIYNRPVSA 533
Db 891 KNOIITHILNISE-----KKLSRVDLLINDLLNNINQINLOISITILMILNNGDINI 947
QY 534 KRIGDIVSVNCI--VVDOTS--VSLHKS-----LRLLSASDEKCFSPPTTFKEMNDST 584
Db 948 KNI--LIKLIINCVKPLFNHTNKDISFHSKIVILLIFLHPKKEYLEPILSFKNILDOT 1005
QY 585 IYKGQL-----GVNN--EILLTTTYLETQOE-----NTEYFQAKTDMYIYKN 625

Db 1006 SYFFLLHISEINNUNVYDILKKNFIDTKEIYLLHMIKEEVNSLQNICYOIFYNLTXYNDN 1065
QY 626 YEHLKTV-PLSGITTTLDTFIALN-----FTLLENVDFKVIETLYTREKRLSNVFD--I 675
Db 1066 NDCLFLYISIIINLLKLPFFIMISIEIHRYYFPAVKNIFFKKNNDIYKENVKPMFMSIMDNL 1125
QY 676 ETMFREYNYAQRVSGRLKDLDLSTNRNQVDFAGSLMDDL 717
Db 1126 LALFNNIPAPKLLGESCTKCVLENN--DENQYKNKLEFLKENM 1166
RESULT 12
PCT-US02-18638A-4
; Sequence 4, Application PC/TUS0218638A
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc. et al.
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY
; TITLE OF INVENTION: OF CERVICAL CANCER
; FILE REFERENCE: MRI-035PC
; CURRENT APPLICATION NUMBER: PCT/US02/18638A
; CURRENT FILING DATE: 2002-06-12
; PRIOR APPLICATION NUMBER: US 60/298,159
; PRIOR FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/298,155
; PRIOR FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/335,936
; PRIOR FILING DATE: 2001-11-14
; NUMBER OF SEQ ID NOS: 238
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 3899
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US02-18638A-4
Query Match 2.8%; Score 126.5; DB 1; Length 3899;
Best Local Similarity 18.7%; Pred. No. 33;
Matches 110; Conservative 101; Mismatches 196; Indels 181; Gaps 26;
QY 373 KEFNTYHCLNEEIKAEIEKKYAKVNSTHSGYDLYKFKTDGGLYLVWQPLIQNRLLDAK 432
Db 605 KEKNAVLDRMAESQEAELERLTQLFSHEE--ELSKLKED-----LEIE 647
QY 433 NKLANNETYSRRSRQAESTTDPMMETGNGAGGEY-----SSENSITVAQVOYAYDNL 485
Db 648 HRIN-----IEKLDNLGIHYKQIDGLQNMESOKIETMOFEKDNL 688
QY 486 RIRINILEDLSKAWCREQ-----HRAALVWNLKSK-----INPT 520
Db 689 ITKQNLILSILSKLDLQOQSLVNSKSEMTLOINELQKEIETLRQEEKEKGTLEQEQVEL 748
QY 521 SVMGMYNRPVSAKRIGDV-----ISVNCIVVDQ-----TSVSLHKSRLLS 563
Db 749 QLKTLELKEQKKE--NDLQEKPAQLEAENSILKDEKTKLEDMLKIHTPVSQEBRLIFLD 807
QY 564 ASDEKCSRPPVPTK-----FNDSTIYKGO--LGVNNIILLTTTYLETCQENTYFQ---- 615
Db 808 SIKSK--SKOSWEKEIEIIEENEDLKQOCIQLNNEEIKORNTFSPAENKFEVNYQELQ 865
QY 616 -----AKTDMYIYKNYEHK-----TVPL--SSITLLD--TFIA 645
Db 866 EYACLLKVKDDLEDSDKXQOEYKSKLKAINEELHLQRIPTTVKXSSVDFDEKTFVA 925
QY 646 LNFTLLENVDFKVIETL-----YTRDEK-----RLSNVFD--IETMFREYNYAQRVSGLRK 694
Db 926 ETLEMGVEVKDTELMEKLEVTREKLELSQRLSDLSQELKQKHGEISFLNEEVKSLKQ 985
QY 695 DL-----LDLSTNRNQVDFAGSLMDDLGAQGVTVNNAVSGVATLFSSTVITGINF 745
Db 986 EKEQVSLRCLRELEIINHNH-----AENVQSDQVSSLLDGVVVTMTSRGAGSVSK 1037
QY 746 IKNPFGGLMIIV-----VIG-----VLFALYFLTKTKIYETAPIK--MIYP 786

Db 1038 VNKSGEESKIMVEDKVSFENMTVGEESKQEQILDLHLPSTKESLRATQPSNDKLOK 1097
Qy 787 EIDKLEREGKSEIAPISEELERIVLAM-----HIHQONSHMETKTRK 830
Db 1098 ELNVKSEQNDLRL-----QMEAQRICLSLVYSTHVDQVREYMEKDK 1141

RESULT 13

PCT-US02-18638A-2
; Sequence 2, Application PC/TUS0218638A
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc. et al.
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY
; TITLE OF INVENTION: OF CERVICAL CANCER
; FILE REFERENCE: MRI-035PC
; CURRENT APPLICATION NUMBER: PCT/US02/18638A
; CURRENT FILING DATE: 2002-06-12
; PRIOR APPLICATION NUMBER: US 60/298,159
; PRIOR FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/298,155
; PRIOR FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/335,936
; PRIOR FILING DATE: 2001-11-14
; NUMBER OF SEQ ID NOS: 238
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 3907
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US02-18638A-2

Query Match 2.8%; Score 126.5; DB 1; Length 3907;
Best Local Similarity 18.7%; Pred. No. 33;
Matches 110; Conservative 101; Mismatches 196; Indels 181; Gaps 26;
Qy 373 KEFTTYVCHLNEEIKAEIEKKYAKVNSTHSGYDLYKFTDGLGYLVWQPLIQNRLLDK 432
Db 605 KEKNAVLDRAESQEAELERLTQLFSHEE--ELSKLKED-----LEIE 647
Qy 433 NKLNNETYSRRSRQAEISTTDPMMMTGNGAGGY-----SSENSITVAQVAYDNL 485
Db 648 HRIN-----IEKLDNLGIHYKQIDGLQNMESQKTIETWQFEXDNL 688
Qy 486 RIRNNILEDLSKAWCREQ-----HRAALVWNLSEK-----INPT 520
Db 689 ITKQNLILEISKLDLQSLVNSKSEMTLQINELQKEIEILRQEEKGTLEQEQVEL 748
Qy 521 SVMSIYNRPVSASRIGDV-----ISVNCIVVDQ-----TSVSLHKSRLLS 563
Db 749 QLTLELLEKQWKE--NDLQEKFAQLEAENSILKDEKTKLEDMLKIHTPVSQBERLIFLD 807
Qy 564 ASDEKCFSPPTPK-----FMNDSTIYKGO--LGVNNEILLTTTYLETQCENTEYFQ--- 615
Db 808 SIKS--SKDSVWEKEIEILIEENEDLKQOCIQLENEIEKQRTNFSFAEKNFVNYQELQ 865
Qy 616 -----AKDMYIKVNYEHLK-----TVPL-SSITLTD-TFIA 645
Db 866 EYACLLKVKDDLEDSKNQKELEYKSLKALNEELHLQINPTTVKMKSSVFDKTFVA 925
Qy 646 LNFTLLENVDFKVEL-----YTRDEK-----RLSNVFD--IETMFREYNYAQRVSLRK 694
Db 926 ETELEGEVVEKDTTELMKELEVTREKLELSQRLSDLSQKQKHGEISFLNEEVSCLKQ 985
Qy 695 DL-----LDLSTNRNQFVDAFGLMDDLGAAGQTVVNAVSGVATLFSIVTGFINF 745
Db 986 EKEQVSLRCRELEIINHN-----AENVQSCDTQVSSLLDGVVVTMTSRGAEGSVSK 1037
Qy 746 IKNPFGMLMIIV-----VIG-----VLFALYFLTKTKIYETAPIK--MIYP 786
Db 1038 VNKSGEESKIMVEDKVSFENMTVGEESKQEQILDLHLPSTKESLRATQPSNDKLOK 1097

Qy 787 EIDKLEREGKSEIAPISEELERIVLAM-----HIHQONSHMETKTRK 830
Db 1098 ELNVKSEQNDLRL-----QMEAQRICLSLVYSTHVDQVREYMEKDK 1141
RESULT 14
PCT-US02-18638A-8
; Sequence 8, Application PC/TUS0218638A
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc. et al.
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY
; TITLE OF INVENTION: OF CERVICAL CANCER
; FILE REFERENCE: MRI-035PC
; CURRENT APPLICATION NUMBER: PCT/US02/18638A
; CURRENT FILING DATE: 2002-06-12
; PRIOR APPLICATION NUMBER: US 60/298,159
; PRIOR FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/298,155
; PRIOR FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/335,936
; PRIOR FILING DATE: 2001-11-14
; NUMBER OF SEQ ID NOS: 238
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 3917
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US02-18638A-8

Query Match 2.8%; Score 126.5; DB 1; Length 3917;
Best Local Similarity 18.7%; Pred. No. 33;
Matches 110; Conservative 101; Mismatches 196; Indels 181; Gaps 26;
Qy 373 KEFTTYVCHLNEEIKAEIEKKYAKVNSTHSGYDLYKFTDGLGYLVWQPLIQNRLLDK 432
Db 605 KEKNAVLDRAESQEAELERLTQLFSHEE--ELSKLKED-----LEIE 647
Qy 433 NKLNNETYSRRSRQAEISTTDPMMMTGNGAGGY-----SSENSITVAQVAYDNL 485
Db 648 HRIN-----IEKLDNLGIHYKQIDGLQNMESQKTIETWQFEXDNL 688
Qy 486 RIRNNILEDLSKAWCREQ-----HRAALVWNLSEK-----INPT 520
Db 689 ITKQNLILEISKLDLQSLVNSKSEMTLQINELQKEIEILRQEEKGTLEQEQVEL 748
Qy 521 SVMSIYNRPVSASRIGDV-----ISVNCIVVDQ-----TSVSLHKSRLLS 563
Db 749 QLTLELLEKQWKE--NDLQEKFAQLEAENSILKDEKTKLEDMLKIHTPVSQBERLIFLD 807
Qy 564 ASDEKCFSPPTPK-----FMNDSTIYKGO--LGVNNEILLTTTYLETQCENTEYFQ--- 615
Db 808 SIKS--SKDSVWEKEIEILIEENEDLKQOCIQLENEIEKQRTNFSFAEKNFVNYQELQ 865
Qy 616 -----AKDMYIKVNYEHLK-----TVPL-SSITLTD-TFIA 645
Db 866 EYACLLKVKDDLEDSKNQKELEYKSLKALNEELHLQINPTTVKMKSSVFDKTFVA 925
Qy 646 LNFTLLENVDFKVEL-----YTRDEK-----RLSNVFD--IETMFREYNYAQRVSLRK 694
Db 926 ETELEGEVVEKDTTELMKELEVTREKLELSQRLSDLSQKQKHGEISFLNEEVSCLKQ 985
Qy 695 DL-----LDLSTNRNQFVDAFGLMDDLGAAGQTVVNAVSGVATLFSIVTGFINF 745
Db 986 EKEQVSLRCRELEIINHN-----AENVQSCDTQVSSLLDGVVVTMTSRGAEGSVSK 1037
Qy 746 IKNPFGMLMIIV-----VIG-----VLFALYFLTKTKIYETAPIK--MIYP 786
Db 1038 VNKSGEESKIMVEDKVSFENMTVGEESKQEQILDLHLPSTKESLRATQPSNDKLOK 1097
Qy 787 EIDKLEREGKSEIAPISEELERIVLAM-----HIHQONSHMETKTRK 830
Db 1098 ELNVKSEQNDLRL-----QMEAQRICLSLVYSTHVDQVREYMEKDK 1141


```

; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114A
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 53825
; LENGTH: 353
; TYPE: prt
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: UC-GMROPIC049A02_Flt.bep
US-10-425-114A-58025

```

```

Query Match      2.7%; Score 124.5; DB 6; Length 353;
Best Local Similarity 20.2%; Pred. No. 4.7;
Matches 80; Conservative 65; Mismatches 143; Indels 109; Gaps 19;

-QY 455 MMEWTGAGGEYSSENSITVAQVQYAYDNLRIIRINILLEDLKAWCREQHQRAALVNNEL 514
      : : : : | | | | | | | | | | : : : : : : : : : : : : : : : : : :
Db 30 LFELLGOSAGDPLVEEDDIPILVRSWQSNFLVTVMHI-----KGSVSW--I 74

QY 515 SKINPTSVMSI-----YNRPSAKRIGDVISVNCIVVDQTSVLSLHKSRLLSASDEKCF 570
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 75 NVLGITEVQELLSAGGYNMP---RTVHEVIALACRLSRWDDRLFRKS--IFGAAD--- 126

QY 571 SRPVTTFKFN-----DSTYKGGOLGVNNEILLTTTILETCQENTEYVFOAKTDMYIKN 625
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 127 ----IELFKPMNRNHEDLNLF--TLINQEI-----RKLSTQIVRKWSLHAR-DEIVPEL 175

QY 626 YEHLKTVPLSSITTLDTFIALNFTLENVDFKVIELYTRDEKRLSNVFDIETMEREYNY 685
      : : : : | | | | | | | | : : : : : : : : : : : : : : : : : :
Db 176 LQHLKNGAR-----TLLEGIKKSTREEMEEQEAVRGLFTIODVWQ----- 217

QY 686 AQRVSGLRKOLLDLSTNRNQFVDAGFSLMDDLGAVGTQVNVNAVSGVATLFSSIVGTGFNF 745
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 218 ----STVRWLQDRSLRVTHNLAVF-----GGVG-VVLITITG---LFGINDVGIPGA 262

QY 746 IKNP--FGGMLMIITWIGLVAIFYLTKTKYITAPIKMYYPEIDKLKEREGKSEIAP I 803
      : : : : | | | | | | | | : : : : : : : : : : : : : : : : : :
Db 263 EQTPYAFGVFTAILVUGVL-----IAGVWYLGKN-----PV 297

QY 804 SEELE-----RIVLAMHIHQONSHMETKTRKDPK 834

Db 298 AEEOVEYRKLEOBLVKMFOHEATHAOMRKNISPKN 334

```

```

RESULT 18
US-10-670-695-16
; Sequence 16. Application US/10670695
; GENERAL INFORMATION:
; APPLICANT: Jensen, Wayne A.
; APPLICANT: Lappin, Michael R.
; APPLICANT: Rosen, David K.
; APPLICANT: Andrews, Janet S.
; TITLE OF INVENTION: USE OF RECOMBINANT ANTIGENS TO DETERMINE THE IMMUNE
; TITLE OF INVENTION: STATUS OF AN ANIMAL
; FILE REFERENCE: DI-9-1
; CURRENT APPLICATION NUMBER: US/10/670,695
; CURRENT FILING DATE: 2003-09-25
; PRIOR APPLICATION NUMBER: 09/521,738
; PRIOR FILING DATE: 2000-03-09
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 16
; LENGTH: 250
; TYPE: Prt
; ORGANISM: Feline herpesvirus 1
; US-10-670-695-16

```

```
Query Match      2.7%; Score 124; DB 6; Length 250;
Best Local Similarity 28.7%; Pred.No.3.7;
Matches 37; Conservative 22; Mismatches 66; Indels 4; Gaps 3;

Qy    80 PSTFPYRVCSAGSDGVDFRFTDTHVCPDAS-DWVHSEGIILLYKONLIIPMFVRVKYRKV 138
       ||||| . : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db    121 PSIF--VMCPPPSGSTVRLPPRPACPDYKLGNFTEGTAVIPEKNIAPIYKKANIYYKN 178

Qy    139 VTTSTVVNGHYGDSITNQHTFYKSIEPWE-TEKMWTIYQCFSNLSRLNTWGNLITYVDRDD 197
        : : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db    179 IIMTWSGSYAVTNNRYTRDRVPKVQEITDLIDRRGMCLSKADYVRNNYQTAFDRDE 238

Qy    198 INMTVFLOP 206
         : : | : |
Db    239 DPRELPLKP 247


RESULT 19
PCT-USO3-27401-450
 ; Sequence 460, Application PC/TUS0327401
 ; GENERAL INFORMATION:
 ; APPLICANT: TUFTS UNIVERSITY
 ; TITLE OF INVENTION: STREPTOCOCCUS PNEUMONIAE FOR DIAGNOSIS, TREATMENT AND PREVENTION OF ACTIVE INFECTION
 ; FILE REFERENCE: 700355-52941-PCT
 ; CURRENT APPLICATION NUMBER: PCT/USO3/27401
 ; CURRENT FILING DATE: 2003-09-02
 ; PRIOR APPLICATION NUMBER: US 60/407,082
 ; PRIOR FILING DATE: 2002-08-30
 ; NUMBER OF SEQ ID NOS: 560
 ; SOFTWARE: PatentIn version 3.2
 ; SEQ ID NO 460
   LENGTH: 1038
   TYPE: PR1
   ORGANISM: Streptococcus pneumoniae
PCT-USO3-27401-450
```

[illegible]

```

Qy   635 SSI-----TTLDTFIALNF-----TLLENVDFKVIELYTRDEKLNSNVF---DIET 677
      |||       :||        |::|::|::|::|::|::|::|::|::|::|::|
Db   621 SSVDYGDNPDRADGNRNGFNFGSVTHTRADNPSWVEVDLKKMDKVGLVKIYNRTDAET 680
      |||       :||        |::|::|::|::|::|::|::|::|::|
Qy   678 MREINYYAQRVSGURKDLLDLSLTNRNQFDAGFSGLMDDLGAVGQTIVNAVSGVATLFSS 737
      |||       :||        |::|::|::|::|::|::|::|::|::|
Db   681 -----QLSNF--DVILYDNRNNE-----VAKHVNLLSGESVSLEDF 715
      |||       :||        |::|::|::|::|::|::|::|::|::|
Qy   738 IVTGFINFIKPFPGGMLIIVVIGVFALFYFLTKTIVETAPIKMVIPEIDKLKEREGK 797
      |||       :||        |::|::|::|::|::|::|::|::|::|
Db   716 KKG-----ARYI---XVKLL-TSGVPLSLAEVVFRESOGK 748
      |||       :||        |::|::|::|::|::|::|::|::|::|
Qy   798 SEIAPISEELERI 811
      |||       :||        |::|::|::|::|::|::|::|::|::|
Db   749 Q-----SEEDIDKI 757
      |||       :||        |::|::|::|::|::|::|::|::|::|

RESULT 20
US-10-472-928-4532
; Sequence 4532, Application US/10472928
; GENERAL INFORMATION:
; APPLICANT: CHIRON SpA
; TITLE OF INVENTION: THE INSTITUTE FOR GENOMIC RESEARCH
; FILE REFERENCE: P026926WO
; CURRENT APPLICATION NUMBER: US/10/472,928
; CURRENT FILING DATE: 2003-09-26
; PRIOR APPLICATION NUMBER: GB-0107658.7
; PRIOR FILING DATE: 2001-03-27
; NUMBER OF SEQ ID NOS: 4979
; SOFTWARE: SeqWin99, version 1.03
; SEQ ID NO 4532
; LENGTH: 1038
; TYPE: PRP
; ORGANISM: Streptococcus pneumoniae
; FEATURE:
; OTHER INFORMATION: fucosactin-related protein
; OTHER INFORMATION: Cellular location: outside
; OTHER INFORMATION: Similar to strain R6 sequence 15904006 (O.E+01)
US-10-472-928-4532
```

		Query Match	2.7%; Score 124; DB 6; Length 1038;
		Best Local Similarity	17.8%; Pred.No.14;
		Matches	120; Conservative 84; Mismatches 216; Indels 254; Gaps 27;
Qy	209	GVTPTDKRYGSOPELYLEPGWFGWSYRRRTVNCELMDMFARSNPPDF----	FVTATGD 264
Db	267	GRADMKSYASEP-----SWIAEMMNVTGGGTVNFECAATFTTND 311	
Qy	265	-----TVEMSPFWSGEDDHENKHEKPFWFVSVINNYKVVDYQNRGTVPLGKTFRFLDRE 318	
Db	312	VPTPAFTKGIIFFRHAIQNPAPSKEE-----VVNRTKAVFWNEGEGR----- 354	
Qy	319	EYTLSWEKHLKNMYSYCPLTLWKAFYNGIQTE-----HSGSYHYFV-----ANDITA 363	
Db	355	-----SSLNGFYQGLYSNDEMTPLYNNRGXYHILPVIHKEIDKEIKSS 396	
Qy	364	SFTTSKEDMKREFNTTYHCLMBEIEKAEIKYAKVNSTHSKYGDLUKPYKTDGGLYLVMQPL 423	
Db	397	IFPNAKILTGN-----SEELS-----SKVNLNSLP--KLVEGDG----- 430	
Qy	424	IONELLDAKNLNNETYSRRSRROAESTTDPDMEMWTGNAGGEYSSENSITVAOVQAYD 483	
Db	431	-----YAQRVGNSWIYNSANINKNQOVMPLPMYTNNTKSLDLTPHTYAVVKNPN 483	
Qy	484	NLRIRINNILEDLSKAWCREQHRAALVYNELSKINPTSVMSMIYNNRPVSAKRIGDVISVS 543	
Db	484	NLHILLNNRYDTKAMW-----ALSNGFDASKSWKK---EEELELA 520	
Qy	544	NCIYVDQTSVLHLKSLRLLSASDEKCPSPRPVTFKFWDNDSTIYQGOLGVNNEILLT----- 599	
Db	521	NWI-----SKNYSINPDNDFRTTTLTUKGHTGHCKPOINISGDN 560	

```

Qy 600 -TTYLETCQENTEY-----FQAKTDMVIYK-----NYEHLKTVPL 634
Db 561 HYTYTENWDNTVTVTVNHNGMVENSINTEGTCPVSPDKFNDGNLNIAYAKPTTQ 620
Qy 635 SSI-----TTLDTFFALNF-----TLLENVDFKVIELYTRDEKLSNVF---DIET 677
Db 621 SSVYNGDPNRAVDGNRNGNSGVSVTHTRADNPSWEVDLKKMDKVGVLKIYNRTDAET 680
Qy 678 MFREYNYYAQRVSGLRKDLLDLSTNRNODVDAFGSLMDDLGAVGOTVVNAVSGVATLFS 737
Db 681 -----QRLSNF--DVILYDNRNE-----VAKHVNLSGESVSLDF 715
Qy 738 IVTGFINKPFGMLMIIVIGVLFAYELTKTKIYETAPIKMIYPEIDKLKEREGK 797
Db 716 KEKG-----ARYI---KVULL-TSGVPLSLAEVEVFRESOGK 748
Qy 798 SEIAPISEELERI 811
Db 749 Q-----SEEDIDKI 757

RESULT 21
US-10-472-928-3638
; sequence 3638, Application US/10472928
; GENERAL INFORMATION:
; APPLICANT: CHIRON SPA
; APPLICANT: THE INSTITUTE FOR GENOMIC RESEARCH
; TITLE OF INVENTION: STREPTOCOCCUS PNEUMONIAE PROTEINS AND NUCLEIC ACIDS
; FILE REFERENCE: P026926WO
; CURRENT APPLICATION NUMBER: US/10/472,928
; CURRENT FILING DATE: 2003-09-26
; PRIOR APPLICATION NUMBER: GB-0107658.7
; PRIOR FILING DATE: 2001-03-27
; NUMBER OF SEQ ID NOS: 4979
; SOFTWARE: SeqWin99, version 1.03
; SEQ ID NO 3638
; LENGTH: 526
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
; FEATURE:
; OTHER INFORMATION: hypothetical protein
; OTHER INFORMATION: Cellular location: cytoplasm
US-10-472-928-3638

```

Query Match	2.7%;	Score 123;	DB 6;	Length 526;	
Best Local Similarity	18.2%;	Pred. No. 8.2;			
Matches	108;	Conservative 107;	Mismatches 252;	Indels 126;	Gaps 27;
Qy	270	PFWSGEDHKNKHE--KPMFVSVINNYKVDVQNRCRTVPLGKTRIF-----LDREYTYLS	323		
Db	6	PAMYGS---ERTWHADITPVYFSHR-----LEPDD---TFHQIRLFQEQDIDSRLLVLVA	54		
Qy	324	WEKHLKNMSPCLTLWKAFYNGIQTEHSGSYHFVANDITASTFTSKEDMKEFNTTYHCLN	383		
Db	55	YQPHRLFYLRHGVLEMDTYSVFDVMQ---DFHNLHTQVLSIRDIEWDDDCFEFYSPFTI-	111		
Qy	384	EEIKABETEKKYAKVNSTHSKYG---DLKYFKTDGGLYLVNQPLIQNRLLLDAKNKLNNETY	440		
Db	112	--IVQNGKGFKAQVE--HGVEGFISDIQYFEPNGQIHM-----HHIVDRDGVFSSIIIF	160		
Qy	441	SRRSRRQAESTTDPMMEMTCN-----GAGGEYSSENSITVAQVQVAYNLRIRINNIL	493		
Db	161	F-----EDGQAYOEYLNKKEGWOFREKLEKGGQVEVNPILGYRPMULTYQNMGDLVAEFF	216		
Qy	494	EDLSKAWCRBQ-----HRAALVMNELSKINPTSMGSMIYNRPVSAKRIGDV-IVSV	543		
Db	217	ENYLQTVVKQDIFMLPSHSHHQLVLDRLPSTNPKLLSLFIGNPOOTFRDLVDVTEKS	276		
Qy	544	NCIIVDQTSVSLHKSRLLSASDEKCFSRPPVTFKFMNDSTIYKGQLGVNNEILLTTTTL	603		
Db	277	DLILVDR-----EDSLRL-----QELPERMHQCYHLLSSFDTRL	311		
Qy	604	ETCOENTYYFOAKTDMYIKYNYEHLKTVPLSSITLDTFIALNFTLLENVDKVI--ELY	662		

Db 312 RLGRS-----QTKESIYFQDFEQIDNQALLOVLSFVA-----ENKDTVEIFCAF 359
Qy 663 TRDEKRLSNVFDIETMPREYNYAQRVSGRLKDLDDLTST--NRNQFVD---AFGSLMDDL 717
Db 360 AASOEOMNEVEGIFVESFTQENIQSENL-GKAIDYGDADENPLEENQHQDLRLQFVNLDEL 418
Qy 718 GAVG-----QTVNAVSGVATLFSSIVTGFPIFNKPPFGMLMIWVI 760
Db 419 DLIKTLEFVRLIVDLNRHPHLYTOIAGISAGIPQINLVETVYVBLKNGY----LLADVT 474
Qy 761 GVLFALYELTKTKIYETAPIMPEIDKLKREGKSEIAPISE--BELERI 811
Db 475 EFSKAAHYTDRLEKNES---LIY-SIDKIKEHTGQOFLKLEKWEIEVKV 523

RESULT 22

US-10-205-516A-22
; Sequence 22, Application US/10205516A
; GENERAL INFORMATION:
; APPLICANT: Zhong, Jun
; TITLE OF INVENTION: A Method For Producing Biologically Active Botulinum Neurotoxins
; FILE REFERENCE: J2btx1
; CURRENT APPLICATION NUMBER: US/10/205,516A
; CURRENT FILING DATE: 2002-07-25
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: Patent in version 3.2
; SEQ ID NO 22
; LENGTH: 1286
; TYPE: PRT
; ORGANISM: Clostridium botulinum
US-10-205-516A-22

Query Match 2.7%; Score 122.5; DB 6; Length 1286;
Best Local Similarity 18.8%; Pred. No. 20;
Matches 147; Conservative 118; Mismatches 274; Indels 243; Gaps 41;

Qy 64 NENRDSQONKNIYGSSTFFRYVCSAGVCDVFRFQTDHVC-----PDASDMV 112
Db 403 NKGFIENSGQNIERNFA---LQKLSSESVDLFP---TKVCLRLTKIEGRNSRDDSTCI 455
Qy 113 HSEGIILYI---KONIIPFMRVRKRVKVTTS--VYNGIYSDSITNQHTFYKSIPEWT 168
Db 456 KVKNNRILPYVADKDSISQEIFE-----NKIITDETWNQ--YSDKFSIDESILDCQVNP 509
Qy 169 EKMDTIYQCFNSRLNTGGNLLTYVDRDDINMTVFLQFVDPDKRYGSGQPELYLEPG 228
Db 510 EIVDPLLPNVNMEPLNLPGBEIVFYD-----DITKYVD---YLSY 547
Qy 229 WFWGYSR-----RRTVNCCLMDMFARSNPPDF---FVATGDTVMSPP--WSGE--D 276
Db 548 YLESQKLSNNVENITLTSVEEALGYSNKIYTFELPSLAEKVNGVQAGLFLNANWEVE 607
Qy 277 DHENKMEKHPWF-----VSVINNY-----KVVDYQNRG-----T 305
Db 608 DFTTNIMKOTLDKISDVSIIPIGPALNIGNSALRGNFNOAFATAGVAFLEGPFEFT 667
Qy 306 VP-LGKTRIFLDREYTLSEKHLKNMSYC---PLTLWKAFYNGIOTEHSGSYHFVANDI 361
Db 668 IPALGVFTFYSSIQER---EKIKTIENCLQORVKRWK-----DSYQWVMSNW 712
Qy 362 TASFTTSKEDMKFN-----TTYHCLN---BEIKAEIEKKYAKV-----N 398
Db 713 LSRITT-----QFNHINYQMYDSLSYQADAIAKAKIDLEYKRYKSGSDKENIKSQVENLKN 766
Qy 399 STHSKYGD-----KYFKTDGGLVWQPLIQNRLDQKNLKNLETYSRRSRQAESTTD 453
Db 767 SLQVKISEAMNNINKFIRECVTY-----LFKMWLPKVIDELNK--FDLRKTLELINLID 819
Qy 454 PMEMTNGAGGEGYSSNSITVAQVAYVNLRI-----NNILEDL 496
Db 820 -r-----SHNIIUVGEV-----DLKAKVNESPENTMPPNIFSYTNNSLLKDI 860

Qy 497 -----SKAWCREQHRALV-----WNELSKINTSVMSMIYNRPVSAKRIGD--VI 540
Db 861 INEYFNSINDSKLSLQKKNALVDTSGYNAEVRGVNDVOLNTIYTNDFKLSSGDKIIV 920
Qy 541 SVSNCI-----VVDQTSVSL-----HKSRLLSASDEK-----CFSRPPVTFKFM 580
Db 921 NLNNILYSAIYENSSVFWIKSKDLTNSHNEVTIINSIEONSGWKLCTRNGNIEWILO 980
Qy 581 NDSTIYKQGLGVANNEILLTTTLETQCENTEYVFOAKTD-----MYIKNYEHUKTVPLS 635
Db 981 DVNRKYSLIFDYSLSHTGY-----TNKWPFTVITNNIMGYMKLYINGELKQSKIE 1034
Qy 636 SITLDTFIALNFTLLENVDFKVIELYTRDEKLSNVFDIETMPREYN--YYAQRVSGLR 693
Db 1035 DLDEVKLDKTIIVFGIDENIQM-LWIRD-----FNFSKELSGNEDINIVYEGQILRNVI 1089
Qy 694 KD 695
Db 1090 KD 1091

RESULT 23

PCT-US02-24459-102
; Sequence 102, Application PC/TUS0224459
; GENERAL INFORMATION:
; APPLICANT: Curagen Corp. et al.
; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: 21402-416C-061 (Cura-716 SMT WO)
; CURRENT APPLICATION NUMBER: PCT/US02/24459
; CURRENT FILING DATE: 2002-08-02
; PRIOR APPLICATION NUMBER: 60/309,501
; PRIOR FILING DATE: 2001-08-02
; PRIOR APPLICATION NUMBER: 60/316,508
; PRIOR FILING DATE: 2001-08-31
; PRIOR APPLICATION NUMBER: 60/354,655
; PRIOR FILING DATE: 2002-02-05
; PRIOR APPLICATION NUMBER: 60/310,291
; PRIOR FILING DATE: 2001-08-03
; PRIOR APPLICATION NUMBER: 60/383,887
; PRIOR FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: 60/310,951
; PRIOR FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: 60/323,936
; PRIOR FILING DATE: 2001-09-21
; PRIOR APPLICATION NUMBER: 60/381,039
; PRIOR FILING DATE: 2002-05-16
; PRIOR APPLICATION NUMBER: 60/311,292
; PRIOR FILING DATE: 2001-08-09
; PRIOR APPLICATION NUMBER: 60/311,979
; PRIOR FILING DATE: 2001-08-13
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 369
; SOFTWARE: CuraSeqList version 0.1
; SEQ ID NO 102
; LENGTH: 1664
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US02-24459-102

Query Match 2.7%; Score 122; DB 1; Length 1664;
Best Local Similarity 18.0%; Pred. No. 27;
Matches 133; Conservative 97; Mismatches 201; Indels 306; Gaps 34;
Qy 151 DSITNOHTFYKSIETPETERKMDTIYQCFNS-LRLNTGGNLLTYVDRDDINMTVFLQFVGD 209
Db 191 DSVHQLHVYHRDIKP-----DNILMDMNGHIRLADFGSCLKLMDEGTQVSSVAVGTFDY 244
Qy 210 VTPDV-----KRYGSOPE-----LYLEPGWF-----WG---SYRRR--- 237
Db 245 ISPEILQAMBEDGKRGYGPCDMMWGLGVCMYEMLYGETPFYAESLVETYGKIMNHRERFQF 304
Qy 238 ----TTVNCCLMDMPAR-----SNPPDFFTATGDTVMSPFWSGEDDHENKMH 283

```
Db 305 PAQVTDVSENADLIRRLCSREHRLGQNGIEDF-----KKHPFFSGIDWDNIRNC 355
Qy 284 EKPWFVSV-----INNYKVVD--YQNRGTVP-----
Db 356 EAPYIPEVSSPTDTSNFDVDDCLKNSETMPPHTAFSGHLLPFVGYTSSCVLSDRS 415
Qy 308 -----LGKTRI-----FLDREYTLSEKHLKNMSYCPITLWKAFYNGIQTEHSGSY 354
Db 416 CLRVTAGPTSLDLVNQVQRTLDNNLATEAYERRIKRLEQEKLESLKQESTQTQVALQY 475
Qy 355 HFVANDITASTTSKEDMKFNTTYHCLNEEIK-----AEIEKKYAKVNSTHSHYK 405
Db 476 STVDGLPTAS-----KOLETKN-----LKEEIEKLRKQVTESSHLEQOOLEANAVROELD 525
Qy 406 DL-----KYFKTDGGLVLMQPLQNR-----LLDAKNKLN----- 437
Db 526 DAFROIKAYEKQIKT-----LQEREDLNKELVQASERLKNQSKELKDAHCQRKL 575
Qy 438 -----ETYSRRSR-----RQAESTTDPMM-----BMT 459
Db 576 AMQEFMEINERLTHTQKQKLARHVRDKKEEVDLMQKVESLRQELRRTERAKKELEVH 635
Qy 460 GNGAGGEYSSENSITVAQVQYAYDNLIRINNILEDL-----SKAWCREQHRAALVWN 512
Db 636 TEALAAEASKDRKLRQESHEYSK-----OLENELEGLKQKQISYSPGVCSEHQ-----Q 685
Qy 513 ELSKINPTSVMSMIYNRPVSARKRIGDVISVNSCIVVDQTSVSLH-----KSLRLLSASDEK 568
Db 686 EITKLTDLKKSIFYEEELSKREG-----IHANEIKNLK----- 720
Qy 569 CFSRPPVTFKFMNDSTIYKQ-QLGVNNEIL-----LTTTYLETQENTYFYFOAKTMYI 622
Db 721 -----KELHDS-----EGQALNKEIMILKDKLEKTRRESQSEREFSEFK----- 763
Qy 623 YKNEYHLKTVPLSSITLDTFIALNFTLLENVDFKVIELYTRDEKRLSNVFD-ITMPRE 681
Db 764 -QOYEREKV-----LQEREDLNKELVQASERLKNQSKELKDAHCQRKL 575
Qy 682 YNYAQRVSGLRKDLDD 698
Db 794 LSIHQOLEEEVVKDLAD 810
```

RESULT 24

```
PCT-US02-24459-100
; Sequence 100, Application PC/TUS0224459
; GENERAL INFORMATION:
; APPLICANT: CuroGen Corp. et al.
; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: 21402-416C-061 (Cura-716 SMT WO)
; CURRENT APPLICATION NUMBER: PCT/US02/24459
; CURRENT FILING DATE: 2002-08-02
; PRIOR APPLICATION NUMBER: 60/309,501
; PRIOR FILING DATE: 2001-08-02
; PRIOR APPLICATION NUMBER: 60/316,508
; PRIOR FILING DATE: 2001-08-31
; PRIOR APPLICATION NUMBER: 60/354,655
; PRIOR FILING DATE: 2002-02-05
; PRIOR APPLICATION NUMBER: 60/310,291
; PRIOR FILING DATE: 2001-08-03
; PRIOR APPLICATION NUMBER: 60/383,887
; PRIOR FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: 60/310,951
; PRIOR FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: 60/323,936
; PRIOR FILING DATE: 2001-09-21
; PRIOR APPLICATION NUMBER: 60/381,039
; PRIOR FILING DATE: 2002-05-16
; PRIOR APPLICATION NUMBER: 60/311,292
; PRIOR FILING DATE: 2001-08-09
; PRIOR APPLICATION NUMBER: 60/311,979
; PRIOR FILING DATE: 2001-08-13
```

```
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 369
; SOFTWARE: CuraSeqList version 0.1
; SEQ ID NO 100
; LENGTH: 1738
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US02-24459-100
```

```
Query Match 2.7%; Score 122; DB 1; Length 1738;
Best Local Similarity 18.0%; Pred. No. 28;
Matches 133; Conservative 97; Mismatches 201; Indels 306; Gaps 34;
Qy 151 DSITNQHTFYKSIIPWETEKMDTYYQCFSN-LRLNTGNNLLTYVDRDDINMTVFLQPVGDG 209
Db 191 DSVHQLHVHRDIK-----DNILMDMNGHRLADFGSCLKLMEDGTIVQSSVAVGTPTY 244
Qy 210 VTPDV-----KRYGQPE-----LYLEPGWF-----WG-----SYRRR--- 237
Db 245 ISPEILQAMEDGKGRYGPECDWMSLGVCMYEMLYGETPFYAESLVETYGKIMNHKRFQF 304
Qy 238 -----TTVNCELMDMFAR-----SNPPDFEFTVATGDTVEMSPWMSGCDHDKMH 283
Db 305 PAQVTDVSENADLIRRLCSREHRLGQNGIEDF-----KKHPFFSGIDWDNIRNC 355
Qy 284 EKPWFVSV-----INNYKVVD--YQNRGTVP-----
Db 356 EAPYIPEVSSPTDTSNFDVDDCLKNSETMPPHTAFSGHLLPFVGYTSSCVLSDRS 415
Qy 308 -----LGKTRI-----FLDREYTLSEKHLKNMSYCPITLWKAFYNGIQTEHSGSY 354
Db 416 CLRVTAGPTSLDLVNQVQRTLDNNLATEAYERRIKRLEQEKLESLKQESTQTQVALQY 475
Qy 355 HFVANDITASTTSKEDMKFNTTYHCLNEEIK-----AEIEKKYAKVNSTHSHYK 405
Db 476 STVDGLPTAS-----KOLETKN-----LKEEIEKLRKQVTESSHLEQOOLEANAVROELD 525
Qy 406 DL-----KYFKTDGGLVLMQPLQNR-----LLDAKNKLN----- 437
Db 526 DAFROIKAYEKQIKT-----LQEREDLNKELVQASERLKNQSKELKDAHCQRKL 575
Qy 438 -----ETYSRRSR-----RQAESTTDPMM-----BMT 459
Db 576 AMQEFMEINERLTHTQKQKLARHVRDKKEEVDLMQKVESLRQELRRTERAKKELEVH 635
Qy 460 GNGAGGEYSSENSITVAQVQYAYDNLIRINNILEDL-----SKAWCREQHRAALVWN 512
Db 636 TEALAAEASKDRKLRQESHEYSK-----OLENELEGLKQKQISYSPGVCSEHQ-----Q 685
Qy 513 ELSKINPTSVMSMIYNRPVSARKRIGDVISVNSCIVVDQTSVSLH-----KSLRLLSASDEK 568
Db 686 EITKLTDLKKSIFYEEELSKREG-----IHANEIKNLK----- 720
Qy 569 CFSRPPVTFKFMNDSTIYKQ-QLGVNNEIL-----LTTTYLETQENTYFYFOAKTMYI 622
Db 721 -----KELHDS-----EGQALNKEIMILKDKLEKTRRESQSEREFSEFK----- 763
Qy 623 YKNEYHLKTVPLSSITLDTFIALNFTLLENVDFKVIELYTRDEKRLSNVFD-ITMPRE 681
Db 764 -QOYEREKV-----LQEREDLNKELVQASERLKNQSKELKDAHCQRKL 575
Qy 682 YNYAQRVSGLRKDLDD 698
Db 794 LSIHQOLEEEVVKDLAD 810
```

RESULT 25

```
US-10-679-063-11928
; Sequence 11928, Application US/10679063
; GENERAL INFORMATION:
; APPLICANT: Edgerton, Michael D
; TITLE OF INVENTION: TRANSGENIC PLANTS WITH IMPROVED PHENOTYPES
; FILE REFERENCE: 38-15(52054)B
```



```
QY 819 QONSHMETKTRKDPKDSIL 837
Db 1086 --NNHINLANTRKKKKGNVL 1102

RESULT 27
PCT-US03-27401-364
; Sequence 364, Application pc/TUS0327401
; GENERAL INFORMATION:
; APPLICANT: TUFTS UNIVERSITY
; TITLE OF INVENTION: STREPTOCOCCUS PNEUMONIAE FOR DIAGNOSIS, TREATMENT AND
; PREVENTION OF ACTIVE INFECTION
; FILE REFERENCE: 700355-52941-PCT
; CURRENT APPLICATION NUMBER: PCT/US03/27401
; CURRENT FILING DATE: 2003-09-02
; PRIOR APPLICATION NUMBER: US 60/407,082
; PRIOR FILING DATE: 2002-08-30
; NUMBER OF SEQ ID NOS: 560
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 364
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
PCT-US03-27401-364

Query Match 2.7%; Score 121; DB 1; Length 2004;
Best Local Similarity 18.5%; Pred. No. 36;
Matches 165; Conservative 108; Mismatches 295; Indels 322; Gaps 43;

QY 28 AETGVTSPNTATWSTESPLTGHYTHDSHGERGNNENRDSE--QNKNIYGSPTFPY 85
Db 669 SENGQTEPEPSNGNSTED-----VSTESNTSNGNEEIKQENELDPDKKVEPEKTLLEL 723

QY 86 RVCASAGVDVFRFQTDHVCPCDASDMVHSEGILLIYKQNI-----IP-----FMRVRKYR 136
Db 724 R-----NVSDLELYSL-----SNG---TYKHISLEQVPSNPNSFYVKVK--- 760

QY 137 KVVTTSTVYNGIY-----SDSITNQHTFYKSIPEWTEKMDTIYQCFNSLRNLNTGNLL 190
Db 761 -----SSSFQDVLPVASISEERKNDKILYKI-----TAKVEKLOQ---EIESRYKDNFT 807

QY 191 TYVDR-----DDINMTVFLQPDVGVTPDKRYGSOPELYLEPQFWGSRRTTVCNCELM 246
Db 808 FYLAKKGTBEETNFTSF-----SNLVKAINQNPS-----GTVHLAASLANANEVE 851

QY 247 M-----FARSNPPDFEFTVATGDTVE-----MSPFWSGEDD- 277
Db 852 LGPDRSYIKDTFTGRLLIGEKDGKNYAIYNLKKPLFENLSGATVEKLSKNVAISGKDDI 911

QY 278 -----HKNMKEKPFVSVINNYKVVDYQN 302
Db 912 GSLANEQNNTKIKQVHVDGVLAGEGIGGLLAKAEQSSITESSEFKGRIINTYETTAAYN 971

QY 303 RG-----TKTRIFLDRE-----EYTLISWE 325
Db 972 ICGMVGHLTGDKALLTKSKATVAISSNTNTSDQTVGGLAGLVDRDAQIDSVAEGDINN 1031

QY 326 KHLKMSYCPLTLKAFYNGIQTEHSGS-----YHFVANDITASETT 367
Db 1032 KHFGRVAGVAGNLWDRDTSQDVR--HAGSLTNVLSDVNVNNGNAITGYHY--NEMKVKDTF 1087

QY 368 SKEDMKFEFTTYVCHLNEEI-----KAEIEKKYAKVN-----STHSK 403
Db 1088 SSKANRVYNT--LVKDEVVSKESPEERCTMLDASQIAKAEINPLILPTVEPLUSTSGK 1145

QY 404 -----YCDLKFKYTDGGL-YLVWQPLI-----QNRLLDAKNKL----- 435
Db 1146 KDSDFSQVAYQAKNLTYNKIEKLLPFYNKATIVKYGNLVNNSLLYQKELLSAVMMKD 1205

QY 436 -----NNETYSRRSRQAEISTDPMMETGNCAG--GEYSSENS-ITVAQVQYAYD 483
Db 1206 NQVITDIVSNKQATANKLLHLYKDDILSEKLDLKYQNDFAKLAEYLSGNTGLLTPNQFLYD 1265
```

```
QY 484 NLRIRINNILEDLSKAWCREQHRALVWNELSKINPTSVMSIYNRPVSAKRIGDVISVS 543
Db 1266 QTSI-IKQVLPDLQKV---DYHSEAI--RKTGLGISPNVKQTELYLEQDFAK----- 1310

QY 544 NCIVVDOTSLSLHKSRLLSASDEKCFSPRPVTFKFNNDSTIYKGLGVNNE-ILLITTY 602
Db 1311 -----TKQLEDLSLKKLLSADAGLASANPVTGELYVD-----KIKRKEALLGLTY 1357

QY 603 LETQCENTEYFQAKTDMYIYKNYEHKTVPLSSITLDTFIAL-----NFTLLENVDF 656
Db 1358 LERMY-NFSYQVNVKDLVLY---HLDFFCKGNASPLDTLIELGKSGFNLLAKNNVDT 1412

QY 657 KVIELYTRDEKRLNVDIETMPREYNYYAQRVSGLRKOLLDLSTNRNQF 706
Db 1413 YGIS-----LASQHTTDLFTLEHY-----RKVFLPNTSNDWF 1447

RESULT 28
US-10-472-928-2306
; Sequence 2306, Application US/10472928
; GENERAL INFORMATION:
; APPLICANT: CHIRON Spa
; TITLE OF INVENTION: THE INSTITUTE FOR GENOMIC RESEARCH
; TITLE OF INVENTION: STREPTOCOCCUS PNEUMONIAE PROTEINS AND NUCLEIC ACIDS
; FILE REFERENCE: P026926WO
; CURRENT APPLICATION NUMBER: US/10/472,928
; CURRENT FILING DATE: 2003-09-26
; PRIOR APPLICATION NUMBER: GB-0107658.7
; PRIOR FILING DATE: 2001-03-27
; NUMBER OF SEQ ID NOS: 4979
; SOFTWARE: SeqMan99, version 1.03
; SEQ ID NO 2306
; LENGTH: 2004
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
; FEATURE:
; OTHER INFORMATION: immunoglobulin A1 protease (iga)
; OTHER INFORMATION: Cellular location: Peptidoglycan-bound (LPXTG)
; OTHER INFORMATION: Similar to strain R6 sequence 15903086 (O.E+01)
US-10-472-928-2306

Query Match 2.7%; Score 121; DB 6; Length 2004;
Best Local Similarity 18.5%; Pred. No. 36;
Matches 165; Conservative 108; Mismatches 295; Indels 322; Gaps 43;

QY 28 AETGVTSPNTATWSTESPLTGHYTHDSHGERGNNENRDSE--QNKNIYGSPTFPY 85
Db 669 SENGQTEPEPSNGNSTED-----VSTESNTSNGNEEIKQENELDPDKKVEPEKTLLEL 723

QY 86 RVCASAGVDVFRFQTDHVCPCDASDMVHSEGILLIYKQNI-----IP-----FMRVRKYR 136
Db 724 R-----NVSDLELYSL-----SNG---TYKHISLEQVPSNPNSFYVKVK--- 760

QY 137 KVVTTSTVYNGIY-----SDSITNQHTFYKSIPEWTEKMDTIYQCFNSLRNLNTGNLL 190
Db 761 -----SSSFQDVLPVASISEERKNDKILYKI-----TAKVEKLOQ---EIESRYKDNFT 807

QY 191 TYVDR-----DDINMTVFLQPDVGVTPDKRYGSOPELYLEPQFWGSRRTTVCNCELM 246
Db 808 FYLAKKGTBEETNFTSF-----SNLVKAINQNPS-----GTVHLAASLANANEVE 851

QY 247 M-----FARSNPPDFEFTVATGDTVE-----MSPFWSGEDD- 277
Db 852 LGPDRSYIKDTFTGRLLIGEKDGKNYAIYNLKKPLFENLSGATVEKLSKNVAISGKDDI 911

QY 278 -----HKNMKEKPFVSVINNYKVVDYQN 302
Db 912 GSLANEQNNTKIKQVHVDGVLAGEGIGGLLAKAEQSSITESSEFKGRIINTYETTAAYN 971

QY 303 RG-----TKTRIFLDRE-----EYTLISWE 325
Db 972 ICGMVGHLTGDKALLTKSKATVAISSNTNTSDQTVGGLAGLVDRDAQIDSVAEGDINN 1031
```

Qy 326 KHLKMSYCLTLWKAFYNGIQTEHSGS-----YHFVANDITASFTT 367
Db 1032 KHFGVAGVAGNLDWRTSGDVR--HAGSLTNVLDVNVNTGNALITGVHY--NEMKVKDTF 1087
Qy 368 SKEDMKBFNTTYHCLNBEI-----KAEIEKKYAKVN-----STHSK 403
Db 1088 SSRANRVYNTV--LVKDEVVSKESFEERGTMLDASQASKABINPLILPTVEPLSTSGK 1145
Qy 404 ----YGLDKYFKTDGGL-VLWQPLI-----QNRLLDAKNL-----435
Db 1146 KDSDFSKVAYYOAKRNLTYKRIEKLFPFYNKATIVKYGNILVNNSILYQKELLSAVMMKD 1205
Qy 436 -----NNEYTSRRSRQASTTDPMMEMTNGAG--GEYSSENS-ITVAQVQVAYD 483
Db 1206 NQVITDIVSNKQANKLLHYKDDLSEKLDLKQNDFAKLAESLGNLTGLTYPNQFLYD 1265
Qy 484 NLRIRINNILEDISKWCROHRAALVWNLKINSKINPTSVMSIYNNRVPVSAKRIGDVISVS 543
Db 1266 QTSI-IKQVLPDLQV--DYHSEAI--RKTGLGISPNVKQTELYLEDOFAK-----1310
Qy 544 NCIVVDOTSVLHKSRLLSASDEKCFSPRPVTFKPMNDSTIYKQGLGVNNE-ILLTTY 602
Db 1311 -----TKOOLEDSLKKJLASDAGLASANPVTGYLYD-----KTKRKEALLGLTY 1357
Qy 603 LETCOENTYFFQAKTDMYIYKXVHLKTVPLSSITLDTFIAL-----NFTLLENVDF 656
Db 1358 LERWY-NFSQOVNVKDLVY----HLDFFGKGNASPLDTLIELGKSGFNNLAKNVDT 1412
Qy 657 KVIELYTRDEKRLSNVFDIETMFREYNYAQRVSGLRKDLDLSTNRNQF 706
Db 1413 YGIS-----LASQHTTDLFTLEHY-----RKVELPNTSNNDF 1447

RESULT 29
US-10-679-063-18007
; Sequence 18007, Application US/10679063
; GENERAL INFORMATION:
; APPLICANT: Edgerton, Michael D
; TITLE OF INVENTION: TRANSGENIC PLANTS WITH IMPROVED PHENOTYPES
; FILE REFERENCE: 38-15(52054)B
; CURRENT APPLICATION NUMBER: US/10/679,063
; CURRENT FILING DATE: 2003-10-02
; PRIOR APPLICATION NUMBER: 60/415,758
; PRIOR FILING DATE: 2002-10-02
; NUMBER OF SEQ ID NOS: 27373
; SEQ ID NO 18007
; LENGTH: 846
; TYPE: PRT
; ORGANISM: Plasmodium yoelii yoelii
US-10-679-063-18007

Query Match 2.7%; Score 120.5; DB 6; Length 846;
Best Local Similarity 16.8%; Pred. No. 17;
Matches 135; Conservative 114; Mismatches 265; Indels 289; Gaps 34;
Qy 16 YLYOVAL-YSLSIATGTGTPPNTATWSTES-PLTGHY--GTHDSSHGGERNNRDEE 71
Db 134 YLTQIGLVYLLSKENSIMPEYAKVQNNMLPCLNTWINSNSHSLYGRK-----183
Qy 72 QNKNIYSPSTFPYRVCSAGSGVDVFRFQTDHVCPCDASDMVHSEGILLIYKQNIIPFMR 131
Db 184 -----QIFSSASSIE-----QYIYILDCRCVEFD--CYVFKNIV-----217
Qy 132 VRKRVKVVTSVNGIYSDSITNQHTFYKSIETWETERKMDTITQOC-----FN 179
Db 218 -----VYHGLYGYKLTSILF-----CDTLACMFGFTSPYPIIL 254
Qy 180 SLRL-----NTGGNLLTYVDRDDINMTVLPQVDGVTDPVKRYGSOPELYLEPGFWG 232
Db 255 SLEIHCXKXKXNLIAKILYILGNQLYIPKTKDEINNIPTN-----295
Qy 233 SYRRRTTVNCBELDMFARSNPPFPDFVFTATGDTVEMSGPFWSGEDDHENKMH-EKPVFVS 291

Db 296 -----NCR--NKFVLYKXHFDDNNYT-----SSFYSIFEGUKVNMVYDELGYFSOI 337
Qy 292 IN-----NYKVDYQNRGTVPGLGKTRIFLDREEEYTLSEWHLKN 330
Db 338 INEQEDIDVDEPKYQIIEKKNEKYNKISIEQNOMPEDEPVNNYIE-----QNNSKN 389
Qy 331 MSYCPULLWKAFYNGIQTEHSGSYHFVANDITASFTTSKEDMKBFNTTY-----HCLNE 384
Db 390 TSYKNNQVQKTNKNSDQ-----NNLCNGENVSQDRDREHVDNSYSSSSSSNCLYE 438
Qy 385 EIKAEIEKKYAKVNSTHSGYKGLKYPKTDGGLYLVWQPLIONRLLDLAKNKLNNETYSRRS 444
Db 439 N-----NTIHKTIQIDEKYRND-----NCSN--NNSGFNNS 468
Qy 445 RROAESTTDPMMEMTNGAGGEYSSENSITVAQVAYDNLRIRINNILEDLS--KAWCR 502
Db 469 VPNTWTSDSI-----TNGCKNMSETAHMSKEBYKKQDNEKYK--NNILKNCCLKGYPV 522
Qy 503 EQHRAALVWNLKINSKINPTSVMSI-----YNRPV-----SAKRIGD-----538
Db 523 RNFQENRTYNEICSISENKFIKLIKNEDDVIKYNQKTLTRVYPSGTRLASTWTFNPLIFW 582
Qy 539 ----VISVNCIVVDQTSVSLHLSRLLSASDEKCFSPRPVTFKPMNDSTIYKQGLGVNNE 595
Db 583 SAGIQFVALNYQYNGLSMLLNKGRFLENGGKHSGYILKPELLRF-NKKHDY-NVLSLNIQ 640
Qy 596 IL-----LTTY-----LETQENTYFFQAKTDMYIYKXVHLKT 631
Db 641 ILSLHQLNLLFSIKNYQEKKLKELFKMDMIQRIOTHKKISNKRKKFKYLOKLEREKS 700
Qy 632 VPLSSITLDTFIALNFTLLENVDFKVIELYTRDEKRLSNVFDIETMFREY-----NY 684
Db 701 PULSDIQSD-----SKKKKISN-----ETLLKKFKDNDNDVY 735

Qy 685 YAQRVSGLRKDLDLSTNRNQFV 707
Db 736 IHKKFANIEKEYEDMLTEYKSFL 758

RESULT 30
US-10-679-063-18197
; Sequence 18197, Application US/10679063
; GENERAL INFORMATION:
; APPLICANT: Edgerton, Michael D
; TITLE OF INVENTION: TRANSGENIC PLANTS WITH IMPROVED PHENOTYPES
; FILE REFERENCE: 38-15(52054)B
; CURRENT APPLICATION NUMBER: US/10/679,063
; CURRENT FILING DATE: 2003-10-02
; PRIOR APPLICATION NUMBER: 60/415,758
; PRIOR FILING DATE: 2002-10-02
; NUMBER OF SEQ ID NOS: 27373
; SEQ ID NO 18197
; LENGTH: 793
; TYPE: PRT
; ORGANISM: Plasmodium falciparum
US-10-679-063-18197

Query Match 2.6%; Score 120; DB 6; Length 793;
Best Local Similarity 18.1%; Pred. No. 17;
Matches 106; Conservative 90; Mismatches 237; Indels 152; Gaps 25;
Qy 315 LDREYVTLSEWHLKKNMSYCLTLWK-----AFYNGIQTEHSGSYHFVANDITASFT 366
Db 32 LARKNNMTWSSGSGYGESEITICRENGLIVESQFVENVENPH-----VNDTASISFL 85
Qy 367 TSKEDM--KEFNNTYHCLNEEIKAEIEKKYAKVNSTHSGYKGLKYPKTDGGLYLVWQPLI 424
Db 86 SCSCNIVTRMLNHHNVIVEDIILNEHKDKYANVLFNFPVYNNNE--KVKNGIYEDY--LL 141
Qy 425 QNRLDAA-----KNKLNNETYSRRSRQASTTDPMM 457
Db 142 NSRLLVTVNSTGHVHLNWEHDESSKYLYNDTLTKNEIKNEDIQKMQRYCENYKQVNLN 201


```
Db 236 FSYEPQAGTVRVNYYRTDGNVYKSLWYMGDVKNPSS-----AQWPDGDTF--TA 284
Qy 262 TG-----DTVEMSPFWSGEDDHENKHEKMPFVSVNNYKVVYQNRG 304.
Db 285 TGKGYRIDIPLEAAREFGFLLDDESKQGGDVKIRK-----ENYKFTDLKNH- 332
Qy 305 TVPLGKTRIFL---DREYTLSEKHLKNMSYCLTLWKAFYNGIQTEHSGSVHFVANDI 361
Db 333 -----SOIFLKDDDESITYNPYYVHDIRM-----TGAQHVGTSSI 367
Qy 362 TASFTT---SKED-MKEFNTTYHCLNEEIKAEIEKKYAKVNSTHS-KYGDLYK- 409
Db 368 ESSFSTLVGAKKEDILKHSNITNLGNKVITTDVAIDEAGKVVYSGDFSTKHPYTVSY 427
Qy 410 ----FKTDGGLVYWQPLIQNRLLDAKNLNNETYS-----RRSRQAEST----- 451
Db 428 NSDQFTTKTS---W-----RLKDETSYDGKLGADLKEEGKQVDLTLSWSPS 470
Qy 452 -----TDP-----MMEMTNGAGGE-YSSSENSITVAQVQYAYDNLIRINN- 491
Db 471 ADKVVVVYDKNDDPKVGVTVALEKGERGTWKQTLDSNKLGITDFTGYYYQYQIERQK 530
Qy 492 ---ILEDLSK---AWCREQHR-----AALVMNELSKINPTSVMSMIYRNPVSAGRIGD 538
Db 531 TVLALDPYAKSLAAMNSDDSKIDDAHKVAKAFAVDPAKLG----QDLTYGKIHNFKTRD 587
Qy 539 VISVNCIVVDQTSVSLHKSRLLSASDEKCFSPRPVTFKPMNDSTIYKQGLGVNNEILL 598
Db 588 AV-IYEAHVRDFTS-----DPAIAKDLTKPFGTFEAFIEKLDYLDGLGVTHIQLL 636
Qy 599 TTYLETQCENTEYFFQAKTDMYIYKNYEHKLTVPPLSSITLDTFIALNFTLLENVDVKV 658
Db 637 PV-----LSYFYFNE-----LKNHERLSYASSNSYNNWGYDPQNY-----FSL 675
Qy 701 TNRNQFVDAFGSLMDDLG 718
Db 727 PNYHFMDADGTPRTSFG 744
```

RESULT 33

```
US-10-472-928-364
; Sequence 364, Application US/10472928
; GENERAL INFORMATION:
; APPLICANT: CHIRON SpA
; APPLICANT: THE INSTITUTE FOR GENOMIC RESEARCH
; TITLE OF INVENTION: STREPTOCOCCUS PNEUMONIAE PROTEINS AND NUCLEIC ACIDS
; FILE REFERENCE: P026926W0
; CURRENT APPLICATION NUMBER: US/10/472.928
; CURRENT FILING DATE: 2003-09-26
; PRIOR APPLICATION NUMBER: GB-0107658.7
; PRIOR FILING DATE: 2001-03-27
; NUMBER OF SEQ ID NOS: 4979
; SOFTWARE: SeqWin99, version 1.03
; SEQ ID NO 364
; LENGTH: 1280
; TYPE: PR1
; ORGANISM: Streptococcus pneumoniae
; FEATURE:
; OTHER INFORMATION: alkaline amylopullulanase, putative
; OTHER INFORMATION: Cellular location: Peptidoglycan-bound (LPXTG)
; OTHER INFORMATION: Similar to strain R6 sequence 15902291 (0.E+01)
```

US-10-472-928-364

```
Query Match          2.6%; Score 118.5; DB 6; Length 1280;
Best Local Similarity 20.3%; Pred. No. 32;
Matches 174; Conservative 102; Mismatches 297; Indels 285; Gaps 47;

Qy 9 QSVL--ALWLYQVALYSLSAETGVTS--PPNTATWSTESPLTGHYGHGTHSSHGGRGN 64
```

```
Db 24 GSVILGASLVLLAMATPTTSSDSESTPTTNEPNNRNTTTLAQPLT-----DTAAGS-GKN 76
Qy 65 ENRDSEQKN-----IYGSPSTFPYRVCSASGVDFRQFQTDHVCPSDASDMVHSEGILL 119
Db 77 ESDISSPGNANASLEKTEEPKPAAPAPQGTG-QDRSSEPTTSTSPVTETKAEPIED 135
Qy 120 IYKQNIIPMERVRKRVKVVTTSTVYNGIYS-DSITNQHTFYKSI1EPW-----ETEK 170
Db 136 NY-----FR1H-VKLUPEENKDAQGLWTWDVE-----KPSNWNPGALSFDKARK 180
Qy 171 MDTIYQCFNSRL-----NTGNNLIT-----YVDRDDINMTV 202
Db 181 DD--YGYLDVVKLGEOAKKISPLINNTAGKNLTGDKSEKLVKPKNEAWLQDQ--YKV 235
Qy 203 F-LQPVGDVTPDKRVYGSQPELYLEPCWFGSVYRRRTTVNCELMDFARNPFPDFVTA 261
Db 236 FSYEPQAGTVRVNYYRTDGNVYKSLWYMGDVKNPSS-----AQWPDGDTF--TA 284
Qy 262 TG-----DTVEMSPFWSGEDDHENKHEKMPFVSVNNYKVVYQNRG 304
Db 285 TGKGYRIDIPLEAAREFGFLLDDESKQGGDVKIRK-----ENYKFTDLKNH- 332
Qy 305 TVPLGKTRIFL---DREYTLSEKHLKNMSYCLTLWKAFYNGIQTEHSGSVHFVANDI 361
Db 333 -----SOIFLKDDDESITYNPYYVHDIRM-----TGAQHVGTSSI 367
Qy 362 TASFTT---SKED-MKEFNTTYHCLNEEIKAEIEKKYAKVNSTHS-KYGDLYK- 409
Db 368 ESSFSTLVGAKKEDILKHSNITNLGNKVITTDVAIDEAGKVVYSGDFSTKHPYTVSY 427
Qy 410 ----FKTDGGLVYWQPLIQNRLLDAKNLNNETYS-----RRSRQAEST----- 451
Db 428 NSDQFTTKTS---W-----RLKDETSYDGKLGADLKEEGKQVDLTLSWSPS 470
Qy 452 -----TDP-----MMEMTNGAGGE-YSSSENSITVAQVQYAYDNLIRINN- 491
Db 471 ADKVVVVYDKNDDPKVGVTVALEKGERGTWKQTLDSNKLGITDFTGYYYQYQIERQK 530
Qy 492 ---ILEDLSK---AWCREQHR-----AALVMNELSKINPTSVMSMIYRNPVSAGRIGD 538
Db 531 TVLALDPYAKSLAAMNSDDSKIDDAHKVAKAFAVDPAKLG----QDLTYGKIHNFKTRD 587
Qy 539 VISVNCIVVDQTSVSLHKSRLLSASDEKCFSPRPVTFKPMNDSTIYKQGLGVNNEILL 598
Db 588 AV-IYEAHVRDFTS-----DPAIAKDLTKPFGTFEAFIEKLDYLDGLGVTHIQLL 636
Qy 599 TTYLETQCENTEYFFQAKTDMYIYKNYEHKLTVPPLSSITLDTFIALNFTLLENVDVKV 658
Db 637 PV-----LSYFYFNE-----LKNHERLSYASSNSYNNWGYDPQNY-----FSL 675
Qy 659 IELYTRD-----EKRLSNVFDIETMFE-----YNYAQRVSGLRKOLL-DLS 700
Db 676 TGMYSDDPKNPEKRIA---EFKNLINEIHKRGMGAILDVVYNHTA-----KVDLPEDLE 726
Qy 701 TNRNQFVDAFGSLMDDLG 718
Db 727 PNYHFMDADGTPRTSFG 744
```

RESULT 34

```
PCT-US03-30720-1209
; Sequence 1209, Application PC/TUS0330720
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Wang, Zhiwei
; APPLICANT: Zhou, Ping
; APPLICANT: Zhang, Jie
; APPLICANT: Ghosh, Malabika
; APPLICANT: Ren, Feiyan
; APPLICANT: Asundi, Vinod
; APPLICANT: Zhao, Qing A.
```


QY 736 SSIV--TGFINFKNPPGMLMIWIGVLFVFKTKTIYETAPIKMIYPEIDKLKE 793
Db 715 DKVETESKLOSHNN-----THEAQMSIEDFQWYEEQSKSEE 754
QY 794 RECKSEIAP-----SEELERIVLA-----MHIHQNSHME--TKTRKOPKDS 835
Db 755 QKLADISSLVSKHITRQRELGVRLSSLGDAARGNKAFLDEHTSAMFVTKDAKRKWT 814
QY 836 ILTRAQNMILKRSGYNLKNASVEML 862
Db 815 FAEQAENDCKAGSSSAVKHCRMETML 841
RESULT 36
US-10-330-773-956
; Sequence 956, Application US/10330773
; GENERAL INFORMATION:
; APPLICANT: David W. Morris
; APPLICANT: Marc Malandro
; TITLE OF INVENTION: Novel Compositions and Methods in Cancer
; FILE REFERENCE: 52945201300
; CURRENT APPLICATION NUMBER: US/10/330,773
; CURRENT FILING DATE: 2002-12-27
; NUMBER OF SEQ ID NOS: 981
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 956
; LENGTH: 880
; TYPE: PRT
; ORGANISM: Mus musculus
US-10-330-773-956

Query Match 2.6%; Score 117; DB 6; Length 880;
Best Local Similarity 17.5%; Pred. No. 28;
Matches 113; Conservative 105; Mismatches 239; Indels 190; Gaps 24;
QY 213 DVKRYGQPELYLEPGWFGSRRRTTNCNCELMDFARSNPPDFVATGDTVMSPFW 272
Db 362 EAERHSSPKRELPFG-----LGDNAEL---M 385
QY 273 SGEDDHENKMKHPWFVSVIN-NYKVVDYQNRGTVPGLKTRIFLDBREYTLSEKHLKNM 331
Db 386 GGQDQMDERDVRFRQLKAELNSVIRKLEDRNTLLADERNELKRSRET---EVOLK-- 440
QY 332 SYCPLTLWKAFYNGIQTEHSGSVHFVANDITASFTTSKEDMKFENTTYHCLNEIEKAE 391
Db 441 ---PL-----VEKKNRMKNKEELLSIQRMEEKLSLTRENVMEMKEKLSAQAS 486
QY 392 KYAKVNSTHSGYKGLKYFKTDGGL-YLVMOPLIQNRLDLAKNKNMETYSRRSRQAES 450
Db 487 LK-----RHTSLNDLSLTDEQEIEFLRLQVLEQQHVIDDLS-LEBERLLRSKRHRGKS 539
QY 451 TTDPMEMTNGAGGEYSSENSITVAQVAYDNLRIRINNILEDLSKAWCREQHRALV 510
Db 540 LKPKKHVETFFGFDSESDSETLSYNTDRTDTPATPEEDLDETTREE--ADLR 597
QY 511 WNELSKINPTSVSMIYNRPVSAKRGDIVSVNCIVVDQTSVLSLHSLLSASDEKCF 570
Db 598 FCQLTR--EYQALQRAY--ALLOEQVGGTLDAR--EARTREQLQADLLRCQAKIED-- 648
QY 571 SRPPTVTFKWNDSITYKGQGVNNEILLTITYLETQENTEYVFOAKTDMYIVKYNVHLK 630
Db 649 -----LEKLIVEKQI-----YRLEMEEN-----QLKSEMQADKQNEL- 682
QY 631 TVPLSSITLDTFIALNFTLLENVDFKIELYTRDEKRLSNVFDIE-TMPREYNYYAQRV 699
Db 683 -----LEFRVLEERE--RRSPAFNLQITTFPENNSALQL 717
QY 690 -----SGLRKOLLDLSTNRNQVDFAFGSLMDDLG-----AVQTVVNAV 728
Db 718 FCHOEGVKNLKIETGTAULTQKMDLEKEDQFSROKGYLEQLEQLEQLEQLEQLEQLEQ 777
QY 729 SGVATLFSSI-----VTGFINFKNPPGMLMIWIGVLFVLF 769

Db 778 DLEATLYNALQOEPGRRASEALSASQREDLQAQAVEKVRQ-----L 818
QY 770 TKTKTIYETAPIK-----MIYPEIDKLKEREKSEIAPISEEHELRIY 812
Db 819 LRQSRFDSQILRRERMELLQAQOKTIREMEGKLELQRRQLKELEEKV 865
RESULT 37
US-10-679-063-18166
; Sequence 18166, Application US/10679063
; GENERAL INFORMATION:
; APPLICANT: Edgerton, Michael D
; TITLE OF INVENTION: TRANSGENIC PLANTS WITH IMPROVED PHENOTYPES
; FILE REFERENCE: 38-15(52054)B
; CURRENT APPLICATION NUMBER: US/10/679,063
; CURRENT FILING DATE: 2003-10-02
; PRIOR APPLICATION NUMBER: 60/415,758
; PRIOR FILING DATE: 2002-10-02
; NUMBER OF SEQ ID NOS: 27373
; SEQ ID NO 18166
; LENGTH: 889
; TYPE: PRT
; ORGANISM: Plasmodium falciparum 3D7
US-10-679-063-18166

Query Match 2.6%; Score 117; DB 6; Length 889;
Best Local Similarity 19.5%; Pred. No. 28;
Matches 151; Conservative 118; Mismatches 303; Indels 204; Gaps 35;
QY 185 TGGNLLTYVD-----RDDINMTVFLQPVGDVTPDVKRYG-----SQPELYLEPGWFGS 233
Db 99 TGGGSLSLDDLYILDURKEQKYTWMTVTKGVTGP-RRIGHVMVYSKPNLIV----FGGN 153
QY 234 YRRRTVNCCELMDFARSNPPDFVATGDTVMSP-----FWS-----G 274
Db 154 DGQNTLNDVVMHV---EMPPFEWVRVILIPNTCKVPPQVRVHSADMCKECPASGMIVIFG 210
QY 275 EDDHENKMKHPWFVSVINNYK---VDYQNRGTVPGLK---TRIFLDBREYTLSEKHL 328
Db 211 GRSANCKSLDDTWGLQRHRDGRMDWVEAPIKKGSPPEARYQHTSVFIGSKIFLGG--- 266
QY 329 KMSYCPLTLWKAFYNGIQTEHSGSVHFVANDITASFTTSKEDMKFENTTYHCLNEIEKA 388
Db 267 RNDGCAVPLSTALYN-----TTIEW 288
QY 389 EIEKKYAKVNSTHSGYKGLKY-FKTDGGLYLVMOPLIQN-----RLDLAKNKNNE- 438
Db 289 VTLPSISIKFRHTSWVY---KYTIYTFGGFSSHQTYPTNELECELCFNLLSLNSLSEK 345
QY 439 --TYRRSRROAESTTDPMMWTG---NGAGEYSSENSITVAQVAYDNLRIRINNIL 493
Db 346 KKSIKOSSLKQKLTNDNLKHSDDLNRVNSYNLSQDVINTQOHNISTNN-QPNVSNEL 404
QY 494 EDL--SKAWCREQHRALAVNLSKINPTSVSMIYNRPVSAKRGDIVSVNCIVVDOT 551
Db 405 YDLKNNASICTNLAVPIVPNVQNVNPNVNTHRNMFDTSSNSVFLSRNRPMSNKI 464
QY 552 SVSLHKLRLLSASDEKCFSRPPTVTFKFMNDSTIYKGQGVNNEILLTITYLETQENTE 611
Db 465 RLSAHAHAHQENGSDPAFLVR---KISIDKLEBGR-KINNGVLCPTPNYISEFKNTV 518
QY 612 YFQAKT-----DMYIKNYEHLKTVPLSSITLDTFIALNFTLLENVDFKIELY 662
Db 519 YDKIITLLNPNTQEIQIYNHNSSEIFIPWANISVLCSI-----VIDIF 564
QY 663 TDEKRL-----SNVFDIETMFREYNYYAQRVSGLRKOLL-DLSTNRNQFVD 708
Db 565 KQEDVMVLRAPIKIYGDHGYDLMRMFOLYKCPVEEDLGEKNAIGDIDSDNYFLG 624
QY 709 AFGSLMDDLGVAGQTVVNAVSGVATLFSSTVGTGFINFKNPPGMLMI- ----VWIGV 762
Db 625 DY-----VDRGS-----NSLEVICLLFA-----LKCKYPKQIHLIRGNHEDVAINS 665

```
QY 763 LFAIYFLTKK-----TKIYETAPIKMIYPEIDKL---KEREKGS--BIA 801
Db 666 LYGOBECKRLKEDVTDKDCSWYQINQVFWFLPGAIVE--DKILCVHGIGKSIQIS 723
QY 802 PISEEELERIVLAMIHQONSHMETKTR--KDP--KOSILTTRAQNMRLKRSYCNL 853
Db 724 DIS--QLKRPVVSPQNLNEQKVTDLWSDFTDNDLSILGTIPNDIRPDGTGHI 777

RESULT 38
US-10-205-516A-18
; Sequence 18, Application US/10205516A
; GENERAL INFORMATION:
; APPLICANT: Zhong, Jun
; TITLE OF INVENTION: A Method For Producing Biologically Active Botulinum Neurotoxins
; FILE REFERENCE: J2btx1
; CURRENT APPLICATION NUMBER: US/10/205,516A
; CURRENT FILING DATE: 2002-07-25
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 18
; LENGTH: 1301
; TYPE: PRT
; ORGANISM: Clostridium botulinum
US-10-205-516A-18

Query Match 2.6%; Score 116; DB 6; Length 1301;
Best Local Similarity 20.8%; Pred. No. 45;
Matches 117; Conservative 78; Mismatches 194; Indels 174; Gaps 31;

QY 257 FVVTATGDTVMSPFWS--GEDDH--ENKMHKPEWFSVNNYK--VVDYQNRGTVPGLKTR 312
Db 256 FFMQST-DTIOAEELYTFGGQDPSIISPTDSKSIYDKVLQNGRIVDRINKVLVCISDPN 314
QY 313 IFLDREYTLSEKHLKNNYSYCLTLKAFYNGIOTHESSGSHVFVANDITASFSTSKEDM 372
Db 315 INI-----NIYKNKFKDKYKFV--EDSEGGKYSI---DV 342
QY 373 KEFNTTYHCLN--BEIKAEIEKKYAKVNSTHSGYGLKPKTDGGLVLMVQPLIQNRL 429
Db 343 ESNFKLYKSLMLGFTET--NTAENY-KJKTASYFSDS-----LPPVK---- 382
QY 430 DAKNNLNNTYSRRRROAESTTDPMMEMTNGAGGEYSSENSITVAQVAYDNLIRI 489
Db 383 -IKNLLDNEIY-----TIERGFNISDKNMCKEYRGQNKAINQ--AY----- 421
QY 490 NNILEDLSKAMCREOHRALVWNLKINPTSVMSMIYNRPV--SAKRIGDIVSVNSCIV 547
Db 422 ---EELSK-----BHLA-----VYKIQMCKSVKIEGRVPGI--CID 452
QY 548 VDOTSVSLHKSRLLSASDEKCFSS-----RPPVTFKFMNDSTIYKGLGVNNEILLTT-- 600
Db 453 VDNEN-----LFFIAKNSFSDDLKNERVEYNTQNN-----YIGNDFPINELLIDTDL 501
QY 601 -TYLETQENTYFYFQAKTDMYIKNVEHLTKVPLSSITLDTFTALNFTLLENVDFKVI 659
Db 502 ISKIELPSENTESLTDNFVDPVPEYKQPAIKKVFDENTIYOYLYSQTFFPL-----NIRDI 557
QY 660 ELYTRDEKRL---SNVFDIETMFRYNYAQRV-----SGLRKDOLLDD---LSTNRNQFV 707
Db 558 SLTSSFDALLVSSKVSYSFFSM--DYIKTANKVVEAGLFAGVWKQIVDDFVIEANKSSTM 615
QY 708 DAFGSL-----MDDLGAQGVTVNAVSGVATLFSISIVTGFINFINPFGGMLM 755
Db 616 DKIAISIVPIVIGALNVGDBTAGK-----NPFESAIEIAGSSILLEFI-----PEL 662
QY 756 ITWIGVLFAYIFLTKTKIYET 778
Db 663 LIPVGVFLLESYINDKNKIKT 685

RESULT 39
```

```
PCT-US03-27401-316
; Sequence 316, Application PC/TUS0327401
; GENERAL INFORMATION:
; APPLICANT: TUFTS UNIVERSITY
; TITLE OF INVENTION: STREPTOCOCCUS PNEUMONIAE FOR DIAGNOSIS, TREATMENT AND
; FILE REFERENCE: 700355-52941-PCT
; CURRENT APPLICATION NUMBER: PCT/US03/27401
; CURRENT FILING DATE: 2003-09-02
; PRIOR APPLICATION NUMBER: US 60/407,082
; PRIOR FILING DATE: 2002-08-30
; NUMBER OF SEQ ID NOS: 560
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 316
; LENGTH: 1881
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
PCT-US03-27401-316

Query Match 2.6%; Score 116; DB 1; Length 1881;
Best Local Similarity 19.3%; Pred. No. 63;
Matches 95; Conservative 89; Mismatches 191; Indels 118; Gaps 24;

QY 432 KNKLNNTYSRRRROAESTTD-----PMEMTNGAGGEYSSE-----NSIT 474
Db 393 KSELNNQI-----DKASSVSPDYSTASYNALGPVLET-----AKGVYASEPVKQPEVNSE 444
QY 475 VAQVQVAYDNLRI---RINNILEDLSKAWCREOHRALVWNLKINPTSVMSMIYNRPV 531
Db 445 -NKLTATDALNVKDTLNNTIAD-AKTKVKE-HYSDRSWQNL-QTEVTKAEKVAANTDA 500
QY 532 SAKRIGDIVSVNSCIVVDQTSVLSLHKSRLLSASDEKCFSRPPV---TFKFMNDSTI--- 585
Db 501 KQSEVNEAVEKLTIATIEKLVELS-EKPILTLTSTDKILEREAVAKYTLENQNKTKIKSI 559
QY 586 ----YKQQLGVNNBIL----LTTYLETQENTYFYFQ---AKTDMYIKNVEHLKTVPL 634
Db 560 TAECLKGEEVINTVLTDDKVTETISAAFPKNLYEYKEYTLSTTMIYDRNGEETETLEN 619
QY 635 SSITLTDTFALNFTLLENVDFKIELYTRDEKLSNVFDIETM-FREYNYAQRVSGLR 693
Db 620 QN-----IQDLUKKVELKNIKRTDLIKYENGKETNESLITTPDDKSNYYLKITSNNQ 672
QY 694 KDLDLSTNRNQFVDAGSLMDDLGAQGVQTVNAVSGVATLFSISIVTGFINFINPFGCM 753
Db 673 KTTI-LAVKNIETTVNGTPVKVTAADNLVS-----RTADNKFEERYVHYIEKP----- 722
QY 754 LMIIVIGVLFAYIFLTKTKIYETAPIKMIYPEIDKLKE-----REGKS-----E 799
Db 723 -----KVHE---DNVYVNFKELVEAIONDPSKEYRLGQSMSARN 758
QY 800 IAPISEBELERIVLAMIHQONSHMETKTRKPKDSILTTRA-----QNMLKRSYCN 852
Db 759 VVPNGKSYITKEFTGKLLSSEGKQFAITELEHPLFNVTNATINNVAFFENVEIERSGQDN 818
QY 853 LKNAESVEMNLTL 865
Db 819 I-----ASLANTM 826

RESULT 40
US-10-472-928-1236
; Sequence 1236, Application US/10472928
; GENERAL INFORMATION:
; APPLICANT: CHIRON SpA
; TITLE OF INVENTION: STREPTOCOCCUS PNEUMONIAE PROTEINS AND NUCLEIC ACIDS
; FILE REFERENCE: P026926WO
; CURRENT APPLICATION NUMBER: US/10/472,928
; CURRENT FILING DATE: 2003-09-26
; PRIOR APPLICATION NUMBER: GB-0107658.7
; PRIOR FILING DATE: 2001-03-27
; NUMBER OF SEQ ID NOS: 4979
```

```

; SOFTWARE: SeqWin99, version 1.03
; SEQ ID NO 1236
; LENGTH: 1881
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
; FEATURE:
;   OTHER INFORMATION: zinc metalloprotease ZmpB, putative
;   OTHER INFORMATION: Cellular location: Peptidoglycan-bound (LPXTG)
;   OTHER INFORMATION: Similar to strain R6 sequence 15902625 (0.E+01)
US-10-472-928-1236

Query Match          2.6%; Score 116; DB 6; Length 1881;
Best Local Similarity 19.3%; Pred. No. 63;
Matches 95; Conservative

Qy 432 KKNLNKNTYRRRRQRAESTD-----PMWMTGAGGEYSSE-----NSIT 474
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 393 KSELNNQI-----DKASSVSPTDYSTASYNALGPVLET-----AKGVYASEPVKQPEVNSET 444
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 475 VAQVOYAYDNLRI---RINNILEDLSKAWCREQHRALVWNLKSNPTSVMSMIYRNPV 531
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 445 -NKUKTAIDALNDVKTELNNITAD-AKTQVKE-HYSDRSWQNL-QTEVTKAEKVAANTDA 500
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 532 SAKRIGDIVSYNCIVVDQTSVSLHKSRLLSASDEKCFGRPPV---TFKFMNDSTI--- 585
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 501 KQSEVNEAVEKLTATIEKLVLS-EKPIILTSTDKKILEREAVAKYTLENQNKTKIKSI 559
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 586 ----YKQGLGVNNEL----LTTYLETQOENTYFYFQ---AKTDMYIYKRYHLKTVPL 634
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 560 TAECLKGEEVINTVLVTDKVTETISAAFKNLEYEYKTYLSTTMIYDRNGGEETLEN 619
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 635 SSITLTOTFALNTLENDVDFKVIELYTRDEKLSNVFDIEM-FREYNYAQRVSGLR 693
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 620 QN-----IQDLUKKVELKNKRTDLTKYENGKETNESLITIPDDKSNYLYKITSNNQ 672
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 694 KDLIDLSTNRNQFVDAFGSLMDDLGAQVGTVWNAVSGVATLFSISIVTGFINFKNPFQGM 753
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 673 KTTL-LAVKNLEETVNGTPVYKVTADNLVS-----RTADNKFEEYVHYIEKP----- 722
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 754 LMIIVGVLFPAIYFLTKKIYETAPIKMIYPIEDIKKE-----REGKS-----E 799
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 723 -----KVHE-----DNVYNFKELVEAIQNDPSKEYRLGQSMGARN 758
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 800 IAPISBEELERIVLAMIHQONSHWETKTRKDPKDSILTRA-----QNMKLRKSCYSN 852
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 759 VVPNGKSYITKBFYTKULSSEKGFAPATELHPLFNVITANNVNFENVEIERSQDN 818
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 853 LKNAESVEMLNTL 865
Db : : : :
Qy 819 I-----ASLANTM 826
Db : : : :

RESULT 41
US-10-679-063-18067
; Sequence 18067, Application US/10679063
; GENERAL INFORMATION:
; APPLICANT: Edgerton, Michael D
; TITLE OF INVENTION: TRANSGENIC PLANTS WITH IMPROVED PHENOTYPES
; FILE REFERENCE: 38-15(52054)B
; CURRENT APPLICATION NUMBER: US/10/679,063
; CURRENT FILING DATE: 2003-10-02
; PRIOR APPLICATION NUMBER: 60/415,758
; PRIOR FILING DATE: 2002-10-02
; NUMBER OF SEQ ID NOS: 27373
; SEQ ID NO 18067
; LENGTH: 2497
; TYPE: PRT
; ORGANISM: Plasmodium yoelii yoelii
US-10-679-063-18067

Query Match          2.6%; Score 116; DB 6; Length 2497;
Best Local Similarity 19.2%; Pred. No. 82;
Matches 164; Conservative 124; Mismatches 314; Indels 250; Gaps 50;

```

RESULT 42

```

US-10-472-928-4712
; Sequence 4712, Application US/10472928
; GENERAL INFORMATION:
; APPLICANT: CHIRON SPA
; TITLE OF INVENTION: THE INSTITUTE FOR GENOMIC RESEARCH
; FILE REFERENCE: P026926WO
; CURRENT APPLICATION NUMBER: US/10/472,928
; CURRENT FILING DATE: 2003-09-26
; PRIOR APPLICATION NUMBER: GB-0107658.7

```


Db 1354 --DLV---RCI---RNSLVHKKDI-LRSITQQ--NLPPSKFEPANDCL-----1391
Qy 596 ILLTTTILETCQENTYEFQAKTDMYIKYKYEHLKTVPLSS 636
Db 1392 -----KSEQSVLLYFECKFPDLLFHLVQOPKSDFNS 1423
RESULT 44
PCT-US03-06962-40
; Sequence 40, Application PC/TUS0306962
; GENERAL INFORMATION:
; APPLICANT: The Forsyth Institute
; TITLE OF INVENTION: Immunogenicity of Glucan Binding Protein
; FILE REFERENCE: 25669-018-061
; CURRENT APPLICATION NUMBER: PCT/US03/06962
; CURRENT FILING DATE: 2003-07-01
; PRIOR APPLICATION NUMBER: 10/383,930
; PRIOR FILING DATE: 2003-03-07
; PRIOR APPLICATION NUMBER: 60/402,483
; PRIOR FILING DATE: 2002-08-08
; PRIOR APPLICATION NUMBER: 60/363,209
; PRIOR FILING DATE: 2002-03-07
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 40
; LENGTH: 1518
; TYPE: PRT
; ORGANISM: Streptococcus salivarius
PCT-US03-06962-40
Query Match 2.5%; Score 113.5; DB 1; Length 1518;
Best Local Similarity 18.1%; Pred. No. 71;
Matches 146; Conservative 116; Mismatches 296; Indels 247; Gaps 35;
Qy 135 YRKVVTTSTVYNGIYSISITNQHTFYKSIETP-----WETEKMDTI 174
Db 393 YRLNRATATQGTIDKSIDLDEQS-----DNHMGDFDLDANDVLSNPVQAEQLNQI 447
Qy 175 YQCNSRLNTGGNLLTYVDRDDINMTVFLOPVGCTPDVKRYGSQLPELYLEPCWFGWSY 234
Db 448 HYLNMWSIVMG-----DKDANFDGIRDAVDNVDADM-----LQLYTNYF--- 488
Qy 235 RRRRTTVCCLMDMFARSNPDPFVTATGDTVMSFPWSGDDHEN-----KMKHP 286
Db 489 REYGVNKKSEANALAH-----ISVLEAWSLNDHNDYNDKTDGAALAMENKQ 533
Qy 287 WFSVINNYKVVDYQNRGTVPGLKTRIFLDREBYTSLW-----EKHLKMSYCPPLTWK 340
Db 534 RLALLFSLAKPIKERTPAVSPLYNNTTQDEKTDWINKGSKAYNEDGTVKQSTIGK 593
Qy 341 AFYNGIQTEHSGSYHFV-----ANDITA-----SFTSKEDMKENFTTVHCL 382
Db 594 --YNEKYGDASGNVYFIRAHNNVQDIIAIIKKEINPKSDGFTTIDAEMKQ---APEIY 648
Qy 383 NEETKAEIEKKYAKVN-----STHSKYGDL-----KYFKTDGGLVYWQPL 423
Db 649 NKDMLSS-DRKYTLNNIPAAVAVMLQNMETIRVYGDLYTDGHHYMETKSPYDITVLN 707
Qy 424 IQNRL-----LDAKNKLN---ETYRRRRRQAEESTTDPMMEMTNGAGGEY 467
Db 708 MKSRIKYVSGQAQRSYWLPTDGRWNSDVELY--RTNEVYTSVRYKDIIMTANDTEG-- 763
Qy 468 SSENSITVAQVQAYDNLRIINNILE---DLSKAWCREQRAALVWNLKSKINPTSVMS 524
Db 764 -SKYSRTSGQVTLVANNPKMLDOSAKLVNEMGIHANQYRALIVG----- 809
Qy 525 MIYNRPVSARKIGDIVISVNCIVVDQTSVLHLSRLLSASDEKCFSPPTPKFMDNST 584
Db 810 -----TADGKNFTSDAIAAGYKVTEDNSGVLTFGANDIKGYE-----TFDMSGFVA 858
Qy 585 IYKQGLGVNNEILLTTTILETCQEN-----TEYFQAKTDMYIKYKYEHLKTVPLSSIT 638
Db 859 VWVPVGASDQDIRVASTEAKKSGELTLKATEAY---DSQLIYEGFSNFQITPDGS-- 912

Qy 639 TLDTFIALNFTLLEND-----FKVIELYTR--DEKRLSNVFDIETMPREYNVYQ 687
Db 913 --DPSYTTNRKKAENVDLFPKSCWCVTSFEMAPQVVSADDDGTFLDVSIQ-----NGYA- 961
Qy 688 RVSGRLKDLIDLSTNRNQFVDFARGLMDLDLGAVGQTVVNAVS-----GVATLFS 736
Db 962 -----FADRYDLAMSKNNKYGSKEDLRDALKALHAGIQAIAWDVDPQIYQLPGKEVVTA 1016
Qy 737 SIVTGFINKPFGGMLMIIVIGVLFAIYFUTKKT--KIYET-----APIKIYYPE 787
Db 1017 TRTDG-----AGRKIADAIID--HSLYVANSKSGRDYQAKYGGEGFLAELKAKYPE 1065
Qy 788 I-----DKLKEREGKSE 799
Db 1066 MFKVNMISTGKPIDDSVKLKQWKA 1090
RESULT 45
US-10-679-063-18152
; Sequence 18152, Application US/10679063
; GENERAL INFORMATION:
; APPLICANT: Edgerton, Michael D
; TITLE OF INVENTION: TRANSGENIC PLANTS WITH IMPROVED PHENOTYPES
; FILE REFERENCE: 38-15(52054)B
; CURRENT APPLICATION NUMBER: US/10/679,063
; CURRENT FILING DATE: 2003-10-02
; PRIOR APPLICATION NUMBER: 60/415,758
; PRIOR FILING DATE: 2002-10-02
; NUMBER OF SEQ ID NOS: 27373
; SEQ ID NO 18152
; LENGTH: 1617
; TYPE: PRT
; ORGANISM: Plasmodium falciparum 3D7
US-10-679-063-18152
Query Match 2.5%; Score 113.5; DB 6; Length 1617;
Best Local Similarity 18.0%; Pred. No. 75;
Matches 105; Conservative 84; Mismatches 193; Indels 201; Gaps 25;
Qy 267 EMSPPFWSGE---DDHENKM-----HEKHPWF-----VSVINNYKVVDYQNRGTV 306
Db 511 ELFPFGNGKLYNREENKIEVATFGCNKNWFEKYDVKIENCKYNNYNNENNNDKNKN 570
Qy 307 PL---CKTRIFLDREBYTSLWKEHLKMSYCPPLTWKAFYNGIQTEHSGSYHFVANDITA 363
Db 571 PGFKGLDHIKEEYD---KMYMDGNMYMGKLGIGNISPKIQRSHTD-----NNI-E 620
Qy 364 SFTTSKEDMKFEN-----TTHCLNEEIKAE-----IEKKYAKVNSTHSGYDLY 409
Db 621 NYISDSSENMTYINMNEVRVLPYNNINNNNNNNNNKLIYNQDGLNSIEHTNNISKY 680
Qy 410 FKTGGLYLWQPLIQNRLLDKNNKLNNETYSRRSRQAEESTTDPMMEMTNGAGGEYSS 469
Db 681 TTVD-----YSNDIL---KTGKETYKELTKRKEEICELE-----NSEDKK 718
Qy 470 ENSI-----TVAQVQAYDNLRIINNILE-----ED 495
Db 719 KHNKHKFLDKMKSKKHIEDFFKEDKICERSEKELKIKFTDILHAGAFGHFHRNKG 778
Qy 496 LSKAWCREQRAALVWNLKSKINPTSVMSMIYNRPVSARKIGDIVISVNCIVVDQTSVSL 555
Db 779 LSKSFVSSDVKRNNSFNSINKKKDQLK--YN-----TL 811
Qy 556 HKSRLLSASDEKCFSPRPVTFKPMNDSTIYKQGLGVNNEILLTTTILETCQENTYEFQ 615
Db 812 EKYETKIGAFDESKLG-----NVRNDKTVSYNQIECKNSI-----NYN 849
Qy 616 AKTDMYIYK-----NYEHLKTVPLSSITLDTFTALNFTLLENVDKVIELYTR--DEKRL 669
Db 850 SKTELNDKKNKDFNFKNKWYKMNDEEKIHLKFFNGLDNDKKKELNSDNGIDDKLL 909
Qy 670 SNVFDIETMFREYNVYQVSVGLRKOLL-----DLSTNRNQFVDAFGSL 713

Search completed: October 28, 2003, 15:40:25
Job time : 44 secs

THIS PAGE BLANK (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 28, 2003, 15:27:11 ; Search time 47 Seconds
(without alignments)
1769.913 Million cell updates/sec

Title: US-10-055-364-24
Perfect score: 4547
Sequence: 1 MAGSLKRGSLWLYQV.....KRSGLYSLKNAESVEMLNTL 865

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_76: *
1: pir1: *
2: pir2: *
3: pir3: *
4: pir4: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2489.5	54.8	854	2 T03107	glycoprotein B - a
2	1953.5	43.0	874	2 S58602	glycoprotein B - e
3	1949.5	42.9	808	1 VGBESM	glycoprotein B pre
4	1925	42.3	807	2 T42924	glycoprotein B - a
5	1848.5	40.7	857	1 Q0BE1L	glycoprotein B - h
6	1787	39.3	874	2 S25530	glycoprotein B - b
7	1128	24.8	907	1 VGBETE	glycoprotein B pre
8	1121.5	24.7	906	1 VGBEC1	glycoprotein B pre
9	1084.5	23.9	822	2 T41941	glycoprotein B - h
10	1052.5	23.1	830	2 T43999	glycoprotein B [im
11	1050.5	23.1	830	1 A44047	glycoprotein B pre
12	1041.5	22.9	830	1 B44047	glycoprotein B pre
13	1040.5	22.9	830	2 T41186	probable glycoprot
14	1039	22.9	821	1 VGBE6S	glycoprotein B - h
15	1034	22.7	928	1 VGBEMC	glycoprotein B pre
16	894	19.7	920	1 D48349	glycoprotein B pre
17	876.5	19.3	865	1 VGBEBR	glycoprotein B pre
18	862	19.0	948	2 A56602	glycoprotein B hom
19	857	18.8	932	1 VGBEB2	glycoprotein gi pr
20	845.5	18.6	903	1 VGBEB1	glycoprotein B pre
21	845.5	18.6	903	1 VGBEK1	glycoprotein B pre
22	841	18.5	904	1 VGBEW7	glycoprotein B pre
23	839	18.5	904	1 VGBEB2	glycoprotein B pre
24	837	18.4	904	1 VGBEK2	glycoprotein B pre
25	836	18.4	883	1 VGBEIS	glycoprotein B pre
26	835	18.4	980	1 VGBEC6	glycoprotein B pre
27	833	18.3	985	1 VGBESA	glycoprotein B pre
28	832	18.3	868	1 VGBE31	glycoprotein B - h
29	830	18.3	883	1 VGBE1L	glycoprotein B pre

30	827	18.2	873	2 S26690	glycoprotein B - i
31	824	18.1	980	1 VGBE2H	glycoprotein B pre
32	822	18.1	933	1 B48349	glycoprotein B pre
33	819.5	18.0	919	1 VGBEQH	glycoprotein B pre
34	819.5	18.0	975	2 T42576	probable envelope
35	805	17.7	917	1 VGBEBH	glycoprotein B pre
36	789.5	17.4	943	2 B48474	glycoprotein B - f
37	777.5	17.1	979	2 JH0109	glycoprotein 14 pr
38	775.5	17.1	913	1 VGBEPS	glycoprotein gi pr
39	722	15.9	928	1 VGBEBG	glycoprotein gi pr
40	162	3.6	1524	2 S68553	surface layer prot
41	158.5	3.5	598	2 T28238	ORF MSV077 hypothe
42	154	3.4	2269	2 T28677	rhopty protein -
43	146	3.2	4688	2 F82885	hypothetical prote
44	145.5	3.2	888	2 A38539	p101 protein presu
45	144	3.2	1252	2 B42771	reticulocyte-bind

ALIGNMENTS

RESULT 1
T03107

Glycoprotein B - alcelaphine herpesvirus 1

C:Species: alcelaphine herpesvirus 1

C:Date: 24-Mar-1999 #sequence_revision 24-Mar-1999 #text_change 24-Nov-1999

C:Accession: T03107

R:Ensser, A.; Pflanz, R.; Pleckenstein, B.

J. Virol. 71, 6517-6525, 1997

A:Title: Primary structure of the alcelaphine herpesvirus 1 genome.

A:Reference number: 214840; MUID:97404659; PMID:9261371

A:Accession: T03107

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-854 <ENS>

A:Cross-references: EMBL:AF005370; NID:G2337967; PIDN:AAC58059.1; PID:G2337975

C:Superfamily: herpesvirus glycoprotein B

C:Keywords: glycoprotein

Query Match 54.8%; Score 2489.5; DB 2; Length 854;
Best Local Similarity 57.0%; Pred. No. 5.3e-140;
Matches 471; Conservative 135; Mismatches 173; Indels 47; Gaps 8;

Qy	64	NENRDSSEQN---	KNIVGSPSTFPRVCSAGVGDFRFOTDHVCPPASDMVHSEGIILLI	120
Db	42	SDNIIRQNNATKAGIHSDFSAFPRVCSASNIGDIFRFQFSHPNTPKQKEHGEIILLI	101	
Qy	121	YKQNIIPFMFRVKRYKRVKVTITSTVYNGIYSDSIITNQHTFVKSIIEPWETEKMDTIYQCFSN	180	
Db	102	FKENIVPVFKRYKRVKVTITSTIYNGIYADAVTNQHVFSKSVPIYETRRNDTIYQCYSN	161	
Qy	181	LRLTNGNLLTYVDRDDIINMTVFLQPDGVTPDVKRYGSOBELYLEPGWFGVSRRTTV	240	
Db	162	LDVTVGNNLLVYTDNDGNSMTVDLPQVDGLSNSVRRVHVSQPEIIHAEPGLLGGYRRRTTV	221	
Qy	241	NCELMDMFARSNPFDFTATGTVMSQPM-----SGEDDHENKMKHKEPWFVSU	291	
Db	222	NCEVTETDARAVPPFRFYFITNIGDTIEMSPFWSKAWNETEFSGEPRDT-----LTV	272	
Qy	292	INNYKVDYQNRGVPLGKTRIFLDRREYTLISWEKHLKNMSYCYPLTLWAKFYNGIQTEHS	351	
Db	273	AKDYRVVDYKFRGTQPOGHTRFIVDKEEYTLISWAAQFNNISYCRWAHWKSFDAIKTEHG	332	
Qy	352	GSYHFVANDITASPTTSKEDMKFNTTYHCLNEIEIKAEIEKKYAKVNSTHISKYGLKFK	411	
Db	333	KSLHFVANDITASPYTPTNTQTVLGHVCLNNTIESELKSLRLAKVAKVNDTHSPNGTAQYVL	392	
Qy	412	TDGGLYLWQPLIQNRLLDAKNLN-----NETYSRSRROAESTTDPMMEMTNG	462	
Db	393	TNGGLLLWQPLVQVKLLDAKNLN-----NETYSRSRROAESTTDPMMEMTNG	444	
Qy	463	AGGYSSENSITVAQVQVAYDNLIRINNILEDLSKAWCREOHRALVWVNSLSKINPTSV	522	

445 IDVVYTAESTILLTQIQFAVDTLRAQNNVLEELSRACREQHRASLMNNELSKINPTSV 504
523 MSMIYNRPVSAKRIGDVISVNSCIVVDQTSVSLHKSURLLSA-SDEKCFSRPPVPTKPMN 581
505 MSSYIGRPVSAKRIGDVISVSHCVVDQSDVSLHRSRMKRVPRDRDKTHECYCSRPPVPTKFIN 564
582 DSTIYKQGLGVNNEILLTTTYLETQCEHTYFYQAKTDMYIKYNYEHLKTVPLSSITTLD 641
565 DSHLYKQGLGVNNEILLTTTAVEICHENTHYFQGGNNMYFYKNRYHRVKTMPVGDVATLD 624
642 TFIALNFTLENVDFKVIETYLTREDEKRLSNVFDIETMFREYNYVAORVSGLRKDLLDLST 701
625 TFWVLNLTLENIDFQVIELYSREEKEMSTAFDIETMFREYNYTQEVTCGLRDLTDLAT 684
702 NRNQFVDAFGLMDMDLGAQGVTVNNAVSGVATLFFSSIVTGFINFIKNPFGGLMIIVIG 761
685 NRNQFVDAFGLMDMDLGVGKTVNLNAVSSVATLFFSSIVSGIINFIKNPFGGLMLFGLIAA 744
762 VLPFIYELTKTKIYETAPIKMIYPEIDKLKREGKSEIAPISEEELERITVLAHHYHQON 821
745 VVITVILLNRKAKRFAQNPQMIYPDITKTTSOREELQVDPISKHELDRLNLAHHYHAS 804
822 SHMETK-----TRKDPKDSILTRAQNMLRKSGYSNLKNAESVE 860
805 KPQESKQDEEGSTSGPAD-WLNKAKNVLRRRAGYKPLKRTDSFE 849

RESULT 2
S55602
glycoprotein B - equine herpesvirus 2
C:Species: equine herpesvirus 2
C:Date: 27-Oct-1995 #sequence_revision 03-Nov-1995 #text_change 26-Aug-1999
C:Accession: S55602
R:Telford, E.A.R.; Watson, M.S.; Aird, H.C.; Perry, J.; Davison, A.J.
J. Mol. Biol. 249, 520-528, 1995
A:Title: The DNA sequence of equine herpesvirus 2.
A:Reference number: S55594; MUID:95302501; PMID:7783207
A:Accession: S55602
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-874 <TEL>
A:Cross-references: GB:U20824; NID:G695172; PIDN:AAC13795.1; PID:G695180
A>Note: the nucleotide sequence was submitted to the EMBL Data Library, February 1995
C:Superfamily: herpesvirus glycoprotein B
C:Keywords: glycoprotein

Query Match	43.0%;	Score	1953.5;	DB 2;	Length	874;			
Best Local Similarity	46.1%;	Pred. No.	3.8e-108;						
Matches	403;	Conservative	147;	Mismatches	250;	Indels	75;	Gaps	18;

QY	11	VLA	W	Y	L	Q	V	A	L	S	I	S	T	A	E	T	C	V	T	S	P	P	N	T	A	T	W	S	T	S	P	L	T	G	H	T	D	S	S	H	G	R	N	N	R	D	S	E	70								
DB	10	V	L	C	W	C	V	-	A	A	L	C	Q	V	A	E	V	V	A	-	-	-	-	E	T	T	P	-	-	-	F	A	T	H	-	-	-	-	-	R	E	V	V	A	E	46											
QY	71	E	Q	N	K	N	I	Y	G	S	P	T	F	F	R	V	C	S	A	G	V	-	G	D	V	F	R	F	O	D	H	V	C	P	D	A	S	D	M	W	S	E	G	I	L	L	I	K	O	N	I	P	F	M	129		
DB	47	E	N	P	A	N	P	-	-	-	L	F	R	V	C	A	S	P	T	G	E	I	F	R	F	L	E	S	C	P	N	T	E	K	D	H	I	E	G	A	L	I	Y	K	T	N	I	V	Y	101							
QY	130	F	R	V	K	R	K	V	T	T	S	T	V	Y	G	I	S	D	S	I	T	N	O	H	T	F	Y	K	S	I	E	P	W	E	T	K	M	D	T	I	Y	Q	F	N	S	L	R	L	M	T	G	N	L	189			
DB	102	F	N	V	K	R	K	I	M	T	S	T	I	Y	G	S	E	D	A	I	T	N	O	H	T	R	S	Y	A	V	P	L	E	V	Q	M	D	H	Y	Y	Q	F	S	A	V	Q	N	E	G	H	V	161					
QY	190	L	T	Y	D	R	D	I	N	M	T	V	E	L	O	P	V	G	T	P	D	V	K	R	G	S	O	P	E	L	Y	L	E	P	G	W	F	G	S	Y	R	R	T	T	N	C	E	L	M	D	N	F	A	249			
DB	162	N	T	Y	D	R	G	N	N	E	A	F	K	P	A	D	G	L	T	S	I	T	R	Y	Q	S	O	P	E	V	A	T	P	R	N	L	L	S	Y	T	T	T	T	N	C	E	V	T	E	N	S	A	221				
QY	250	R	S	N	P	P	F	O	F	F	V	T	A	T	G	D	T	E	M	S	P	F	W	S	G	E	D	D	H	E	N	K	M	E	K	P	W	F	S	V	I	N	N	Y	K	V	D	Y	Q	N	R	T	V	P	L	G	309
DB	222	R	S	M	K	P	F	F	F	V	T	S	V	G	D	T	E	M	S	P	L	X	E	N	G	T	E	P	E	K	I	L	K	R	P	H	S	I	Q	L	L	K	N	Y	A	T	K	Y	-	-	-	-	277				
QY	310	K	-	-	-	T	R	I	F	D	R	E	Y	T	L	S	E	K	H	L	N	M	S	Y	C	P	I	T	L	K	A	F	Y	N	G	I	O	T	E	H	S	G	S	Y	H	F	V	A	N	D	I	T	A	S	F	365	
DB	478	Q	A	D	N	A	T	R	F	F	A	I	G	D	S	L	S	W	K	A	T	T	E	N	S	S	C	D	I	L	W	K	F	G	S	N	A	T	O	H	N	S	L	H	I	A	N	D	I	T	A	S	F	337			

Qy	366	TTSKEDMKFENTYHCLNDEEIKAEBIKKAKVNSTHSKYGDLYKFKFTDGGLYLVWQPLIQ	425
Db	338	STPLEEAAFNFTFKIWNNTQEEIOKKLKEVEKTHRPNGTAKVYKTKTGNLYIVWQPLIQ	397
Qy	426	NRLDLAKKLN-----ETYSRRSRQAESTTDPMMEMTCNGAGGEYSS--ENS	472
Db	398	IDLDTLHAKLYNLTNATASPTPTPTSPRRRRDTSS-----VSGGNNGDNSTKEES	450
Qy	473	ITVAQVQAYDNLRIIRINILEDLSKAWCREQRAALVMNELSKINPTSVMSMIYNRPVS	532
Db	451	VAASQVQPAYDNLKRSINRVIGELSEAWCREQYRASLMWYELSKINPTSVMSAIYGRPV	510
Qy	533	AKRIGDVISVSNCIIVVDQTSVLSHLKSLRLLSASDEKCFSPPTVTFKFNWDSTIYKQGQGV	592
Db	511	AKLIGDVVSVDICISVDQKSVFVHKMK--VPGKEDLCYTRPVVGFKPINGSELFAQGLGP	569
Qy	593	NNEILLTTTLYLETCENTYYEYFOAKTDMYIKNYEHLKTVPLSSITLDTFIALNFTLLE	652
Db	570	RNEIVLSTSQVEVQCHSCBHYFQAGNQMYKYDYVYSTLNLTDIPTLHTWITLNLSLVE	629
Qy	653	NVDFKVIELYTRDEKRLSNVFIETMFREYNYYAQRVSGLRKOLL--LSTNRNQFVDAPG	711
Db	630	NIDFKVIELYKTEKRLSNVFIETMFREYNYYTQNLNGLRKOLDSDSIDHGRDSFIOTLG	689
Qy	712	SLMDDLGAQVTVVNAVSGVATLFSVITGVTFNFINKPFCGMLMIIVVI--GVLFATVFLT	770
Db	690	DIMQDLGTIGKVVVNAVSGVFLSGVISGVISVFFKNPFCGMLLIIVLIAGVVVVVYLFMT	749
Qy	771	KTKIYETAPIKMIYPEIDKLKEREKSEIAPISEEELERIVLAMHIHQNSHMETKTRK	830
Db	750	RSRSIY--SAPIRMLPGVERAAQEPGAH---PVSEDIQIRNLIMGHOFQQRQRAEEARR	805
Qy	831	DPK-----DSILTRAQNMLRKR---SGYSNLK	854
Db	806	EEVKVKRGLTPEVIRDSATSVLRRRRGGGGYQRLQ	840
RESULT 3			
VGBESM			
Glycoprotein B precursor - saimirine herpesvirus 1 (strain 11)			
C:Species: saimirine herpesvirus 1			
C:Date: 31-Dec-1991 #sequence_revision 31-Dec-1991 #text_change 16-Jul-1999			
C:Accession: D34126			
R:Abrecht, J.C.; Fleckenstein, B.			
Virology 174, 533-542, 1990			
A:Title: Structural organization of the conserved gene block of Herpesvirus saimirini			
A:Reference number: A34126; MUID:90163221; PMID:2154888			
A:Accession: D34126			
A:Molecule type: DNA			
A:Residues: 1-808 <ALB>			
A:Cross-references: GB:M31122; NID:g331052; PIDN:AAA46164.1; PID:g331056			
C:Superfamily: herpesvirus glycoprotein B			
C:Keywords: glycoprotein; transmembrane protein			
F:1-17/Domain: signal sequence #status predicted <SIG>			
F:18-808/Product: glycoprotein B #status predicted <GLB>			
F:675-692/Domain: transmembrane #status predicted <TML>			
F:702-722/Domain: transmembrane #status predicted <TM2>			
F:730,158,239,251,285,331,344,355,361,471,532,569,587,727/Binding site: carbohydrate			
Query Match 42.9%; Score 1949.5; DB 1; Length 808;			
Best Local Similarity 46.6%; Pred. No. 5.8e+108;			
Matches 373; Conservative 159; Mismatches 236; Indels 33; Gaps 10;			
Qy	71	BQKNKIYSPSTFPYRVCASGVDFRFTQDHVCPDASDMVHSEGIILYKQNIIPFMF	130
Db	28	EKNKTOAIYQYFYKRVCSASTTGLFRFDLDRTCPSTEDKVHKEGILLVYKKNIVPFI	87
Qy	131	RVRYRKVVTTSTVNGIYSD--SITNQHTFYKSIPEWETEKMDTIYQCNSLRNLNTGN	188
Db	88	KVRRYKKTITSVRIENGWTRGVATNKNWELSGRAVPKVIDINDKTYQCHNCQIYNGM	147
Qy	189	LLTVYDRODINNTVFLQPDVGTDPVVKRYGSOPELYLEPGWFVSGYRRTTIVNCELMDMF	248

[illegible]

Db 852 YTRL 855

RESULT 7
VGBETE

glycoprotein B precursor - human cytomegalovirus (strain Towne)
N;Contains: glycoprotein gp55
C;Species: human cytomegalovirus, human herpesvirus 5
A;Note: host Homo sapiens (man)
C;Date: 31-Dec-1989 #sequence_revision 31-Dec-1989 #text_change 16-Jul-1999
C;Accession: A31288

R;Spaete, R.R.; Thayer, R.M.; Probert, W.S.; Masiarz, F.R.; Chamberlain, S.H.; Rasmussen
Virology 167, 207-225, 1988

A;Title: Human cytomegalovirus strain Towne glycoprotein B is processed by proteolytic
A;Reference number: A31288; MUID:89045645; PMID:2460994

A;Accession: A31288

A;Molecule type: DNA

A;Residues: 1-907 <SPA>

A;Cross-references: GB:M22343; NID:g330480; PID:AAA45920.1; PID:g330481

C;Superfamily: herpesvirus glycoprotein B

C;Keywords: glycoprotein; transmembrane protein

F;1-24/Domain: signal sequence #status predicted <SIG>

F;461-907/Product: glycoprotein gp55 #status predicted <GP5>

F;715-748/Domain: transmembrane #status predicted <TM1>

F;751-773/Domain: transmembrane #status predicted <TM2>

F;68,73,85,208,281,286,302,341,383,405,409,417,447,452,456,466,524,555,586,864/Binding

Query Match 24.8%; Score 1128; DB 1; Length 907;
Best Local Similarity 30.9%; Pred. No. 4.2e-59;
Matches 286; Conservative 184; Mismatches 353; Indels 102; Gaps 25;

Qy 14 LWLYQVALYSLSAETG--VTSPPTATWSTESPLTGHYGTDDSS--HGERGNNENRDS 69
Db 5 IWCL--VVCVNLCLVCLGAAVSSSSTRGTSATHS----HHSSHTTSAHSSRSGSVQRVT 58
Qy 70 EQ-----NKNYIG-----SPSTFPRVCSASGVDFPQTDHVC----PDA 108
Db 59 SSQTVSHGVNETIYNTLTLYGVGVVNTTKYPRVCSMAQGTDLIRFERNIVCTSMKPIN 118
Qy 109 SDMVHSEGLLIYKONIIPMFVRVRYKRVVTTSTVYNGIYSDSIINQHTFYKSIPEPWT 168
Db 119 EDL--DEGINVVYKRNIIAHTFKURVQKVLTPFRSAYIHTTYLLGNSTEYVAPPMWEI 176
Qy 169 EKMDTIYQCFNSURLNTGNNLLTYVDRDDI--NMTVFLOQVDPVDPVKRYGSOPELYLEP 227
Db 177 HHINSHQCYSSYSRVIAIGVFVAYHRDSYENKTMQLMPDDYNSHTSTRVTVVKDQWHSR 236
Qy 228 GWFNGSVRRRTTNCCLMDMFAKSNPPDFVTATGDTVMSPFPWGEDDHENKMHKWP 287
Db 237 GSTW-LYRETCLNLCMWITITARSKYPHYFATSTGDDVDISPFGYGTNRNNSYFGENAD 295
Qy 288 FVSVINNYKVVDYQNRGTVPGLKTRI--FLDRREYTLWE-KHLKNMYSYCLTLWKAFYN 344
Db 296 KFIFPNYITVSDFGRNSALETHRLVAFUERADSVISWDIQEKNVT-CQLTFWEASER 354
Qy 345 GIOTEHSGSYHFVANDITASFTTSKEDMKEFNTTYHCLNEEIKAEIEKKY-AKVNQTHSK 403
Db 355 TIRSEADSYPFSKAKMTATFLSKQEVNMSDALOCVRDEAINKLIQIIFNTSYNOTYK 414
Qy 404 YGLKYPKTDGGLVYVWQPLIQNRLLDAKNKLNNETYSRRSRROAESTTDPMMEMTGNGA 463
Db 415 YGNVSVFETGGLVWFVWQGLKQKSLVELERLANRSSLNLTNHRKST-----DGNNA 467
Qy 464 G--GEYSSENSITVAQOVAYDNLRIINNILEDLSKAWCREQHRAALVMNLSKINPTFS 521
Db 468 THLSNMESVANLVYAQLQFTYDLRGYNRALAQIAEAWCQDQRTLEVFKELSKINPSA 527
Qy 522 VMGMINRPSAKRIGDVISVNSCIVVDQTSVSLHKSRLLSASDEKCFSRPPTVKFWN 581
Db 528 ILSAIYKPIAARFMGDVGLGASCVINQTSVKVLRDMN-VKESPGRCYSRPVVIENFAN 586
Qy 582 DSTIYKQGLGVNNEILLTTTLYLETCEBNTYFFQAKTDMYIYKNYEHKLTVPVLSSITLID 641

A:Reference number: Z22732; MUID:99412319; PMID:10492554
A:Accession: T43999
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-830 <ISE>
A:Cross-references: EMBL:AB01506; NID:94995977; PID:BAA78260.1; PID:g4996027
A:Experimental source: strain HST; pop. variant B
C:Genetics:
A:Note: U39
C:Superfamily: herpesvirus glycoprotein B

```
Query Match      23.1%; Score 1052.5; DB 2; Length 830;
Best Local Similarity 32.4%; Pred. No. 1.le-54;
Matches 245; Conservative 157; Mismatches 295; Indels 59; Gaps 15;

Qy      83 PPRVCSAGVDVRFQTDHVC-PDASDMVHSEGILLIYKQNIIPFMRVRKRVKVT 141
Db      36 YPRICSTAKGTDLRMFRDRDISCPYKNAKMSGFFIYKNTIETFPVRYKELTF 95

Qy      142 STVYNGIYSDITNQHTFYKSIETPEWTEKMDTIYQCFNSLRN--TGGNLLTYVDRDIN 199
Db      96 QTSYRDVGYYFLDRTVMGLAMPYEAANLVNSRAQCYSAVAIKRPGDGVFSAYHEDNNKN 155

Qy      200 MTFVLPQVDGVTDPVKRYGQDELYLEPCGFWGSGYRRRTTNCCLMDMFARNSPPDFV 259
Db      156 ETLEFLPLNFKSVTKRFTTKPEYFARGPLW-LYSTSTSLNCIVTEATAKAKYPSYFA 214

Qy      260 TATGDTVMSPFSGED-DHENKMEKHPWFVSVINNYKV--DYON--RGTVPLGKTRIFL 315
Db      215 LTTGIVGSPFPDGSNGKHFAEPLEK---LTILENTMIEDLMNGMGATTLVRKIAFL 271

Qy      316 DREETYLSWEKHLKMSYCPYTLKAFYNGIOTESHSGSYHFVANDITASFTTSKEDMKEF 375
Db      272 EKGDTLFSWEIKEENESVCMKHWTTVTHGLRAETDEIYHFISKELTAFAFVTPKDSINLT 331

Qy      376 NTTYHCLNEEIKAEIEKYY-AKVNSTHSGYGLKVFYKTDGGLYLVWQPLIQNRLDANK 434
Db      332 DPQKTCIKNEFENIIEKVMSYNDYNTYSMNGSYQIFKTTGDLILLIWPQVKSMLF---V 387

Qy      435 LNNETYSSRRRQAESTTDPMMEMTGAGGYSSENSITVAQVQYAYDNLRIINNI 494
Db      388 LEQGSVNLRRRLDV-----DVKSRHDILYVQLQYLDYTLKDYINDALG 431

Qy      495 DLSKAWCREQRAALVWNLKINPTSVMSMIYNRPVSAGRIQDVISVNCIVVDQTSVS 554
Db      432 NLAESWCLDQKRTITMLHLSKISPSISVSEYGRPISAQLHGDVLAISKIEVNOSSVQ 491

Qy      555 LHKSLRLLSA-----SDEKCFSPPTVPKFNDS--TIYKQGLGVNNEILLTYYLETQEN 609
Db      492 LHKSMRVVDKAGLRSETVCYNRPVTFSPVNSTPVPVQGLDNEILLGDHRTCECEIP 551

Qy      610 TEYVFOAKTDMYIYKNYEHKTVPLSSITTLDTPIALNFTLLENVDFKVIELYTRDEKRL 669
Db      552 STKIFLSGNHAYVYDTHNTPTIEDIEVLDAFIRLKIDPLENADFKLLDLYSPDELSR 611

Qy      670 SNVFDIETMFREYNYAQRVSGRLKDLDDLSTNRNQVDFDAFGLSMDLLGAVGQTVNNAVS 729
Db      612 ANVFDLENILREYNSYK---SALYTIKATNTPSYVNGINSFLOGGAGTGTGLGVSIS 668

Qy      730 GVATLFSSIVTGFINFKNPFGGLMIIVIGVLFALYFLTKTKTIYETAPIKMIYP--- 786
Db      669 VTAGALGDIVGGVSWFLKNPFGGLMLILAIIVVVIIIVFVRQKHVLSKPIDMMFPYAT 728

Qy      787 -----EIDKLEREGKSEIAPISEEE 807
Db      729 NPVTTVSSVTGTTAVKTPSVKADGGTSAV-VSEKE 763
```

RESULT 11

A44047

Glycoprotein B precursor - human herpesvirus 6 (strain GS)

C:Species: human herpesvirus 6

C:Date: 17-Feb-1994 #sequence_revision 17-Feb-1994 #text_change 16-Jul-1999

C:Accession: A44047
R:Chou, S.; Marousek, G.I.
Virolgy 191, 523-528, 1992
A:Title: Homology of the envelope glycoprotein B of human herpesvirus-6 and cytomegalovirus
A:Reference number: A44047; MUID:93033157; PMID:1329336
A:Accession: A44047
A:Molecule type: DNA
A:Residues: 1-830 <CHO>
A:Cross-references: GB:M97928; NID:9325474; PID:AAA43847.1; PID:g325475
C:Superfamily: herpesvirus glycoprotein B
C:Keywords: glycoprotein; transmembrane protein
F:1-23/Domain: signal sequence #status predicted <SIG>
F:24-395/Domain: Product: Glycoprotein B #status predicted <GPB>
F:24-395/Domain: gp116 #status predicted <G116>
F:400-830/Domain: gp55 #status predicted <G55>
F:656-674/Domain: transmembrane #status predicted <TM1>
F:693-710/Domain: transmembrane #status predicted <TM2>
F:155,228,247,286,329,355,361,486/Binding site: carbohydate (Asn) (covalent) #status pr

```
Query Match      23.1%; Score 1050.5; DB 1; Length 830;
Best Local Similarity 32.3%; Pred. No. 1.5e-54;
Matches 244; Conservative 160; Mismatches 293; Indels 59; Gaps 15;

Qy      83 PPRVCSAGVDVRFQTDHVC-PDASDMVHSEGILLIYKQNIIPFMRVRKRVKVT 141
Db      36 YPRICSTAKGTDLRMFRDRDISCPYKNAKMSGFFIYKNTIETFPVRYKELTF 95

Qy      142 STVYNGIYSDITNQHTFYKSIETPEWTEKMDTIYQCFNSLRN--TGGNLLTYVDRDIN 199
Db      96 QSSYRDVGYYFLDRTVMGLAMPYEAANLVNSHAQCYSAVAMKRPDGTVFSAFHEDNNKN 155

Qy      200 MTFVLPQVDGVTDPVKRYGQDELYLEPCGFWGSGYRRRTTNCCLMDMFARNSPPDFV 259
Db      156 NTLNLFPLNFKSIYKRPITTKPEYFARGPLW-LYSTSTSLNCIVTEATAKAKYPSYFA 214

Qy      260 TATGDTVMSPFSGED-DHENKMEKHPWFVSVINNYKV--DYON--RGTVPLGKTRIFL 315
Db      215 LTTGEIVGSPFPNGSKHFAEPLEK---LTILENTMIEDLMNGMGATTLVRKIAFL 271

Qy      316 DREETYLSWEKHLKMSYCPYTLKAFYNGIOTESHSGSYHFVANDITASFTTSKEDMKEF 375
Db      272 EKADTLFSWEIKEENESVCMKHWTTVTHGLRAETDEIYHFISKELTAFAVAPKESLNT 331

Qy      376 NTTYHCLNEEIKAEIEKYY-AKVNSTHSGYGLKVFYKTDGGLYLVWQPLIQNRLDANK 434
Db      332 DPQKTCIKNEFENIIEKVMSYNDYNTYSMNGSYQIFKTTGDLILLIWPQVKSMLF--- 388

Qy      435 LNNETYSSRRRQAESTTDPMMEMTGAGGYSSENSITVAQVQYAYDNLRIINNI 494
Db      389 ---EQSEKIRRRRDVV-----DVKSRHDILYVQLQYLDYTLKDYINDALG 431

Qy      495 DLSKAWCREQRAALVWNLKINPTSVMSMIYNRPVSAGRIQDVISVNCIVVDQTSVS 554
Db      432 NLAESWCLDQKRTITMLHLSKISPSISVSEYGRPISAQLHGDVLAISKIEVNOSSVQ 491

Qy      555 LHKSLRLLSA-----SDEKCFSPPTVPKFNDS--TIYKQGLGVNNEILLTYYLETQEN 609
Db      492 LHKSMRVVDKAGRVSETVCYNRPVTFSPVNSTPVPVQGLDNEILLGDHRTCECEIP 551

Qy      610 TEYVFOAKTDMYIYKNYEHKTVPLSSITTLDTPIALNFTLLENVDFKVIELYTRDEKRL 669
Db      552 STKIFLSGNHAYVYDTHNTPTIEDIEVLDAFIRLKIDPLENADFKLLDLYSPDELSR 611

Qy      670 SNVFDIETMFREYNYAQRVSGRLKDLDDLSTNRNQVDFDAFGLSMDLLGAVGQTVNNAVS 729
Db      612 ANVFDLENILREYNSYK---SALYTIKATNTPSYVNGINSFLOGGAGTGTGLGVSIS 668

Qy      730 GVATLFSSIVTGFINFKNPFGGLMIIVIGVLFALYFLTKTKTIYETAPIKMIYP--- 786
Db      669 VTAGALGDIVGGVSWFLKNPFGGLMLILAIIVVVIIIVFVRQKHVLSKPIDMMFPYAT 728

Qy      787 -----EIDKLEREGKSEIAPISEEE 807
```

Db 729 NPVTTSVVTGTTVVKTPTSVKVDGTSVA-VSEKE 763

RESULT 12
B44047
glycoprotein B precursor - human herpesvirus 6 (strain Z29)
C:Species: human herpesvirus 6
C>Date: 17-Feb-1994 #sequence_revision 17-Feb-1994 #text_change 19-Jan-1996
C:Accession: B44047
R:Chou, S.; Marousek, G.I.
Virology 191, 523-528, 1992
A:Title: Homology of the envelope glycoprotein B of human herpesvirus-6 and cytomegalovirus
A:Reference number: A44047; MUID:93033157; PMID:1329336
A:Accession: B44047
A:Molecule type: DNA
A:Residues: 1-830 <CHO>
A:Cross-references: GB:M97927
C:Superfamily: herpesvirus glycoprotein B
C:Keywords: glycoprotein; transmembrane protein
F:1-23/Domain: signal sequence #status predicted <SIG>
F:24-395/400/Product: glycoprotein B #status predicted <GPB>
F:24-395/Domain: gp116 #status predicted <G116>
F:400-830/Domain: gp55 #status predicted <G55>
F:656-674/Domain: transmembrane #status predicted <TM1>
F:693-710/Domain: transmembrane #status predicted <TM2>
F:155,247,286,329,361,486/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 22.9%; Score 1041.5; DB 1; Length 830;
Best Local Similarity 32.3%; Pred. No. 5.1e-54;
Matches 244; Conservative 155; Mismatches 298; Indels 59; Gaps 15;

Qy 83 FPYVCSASGVDFRQTDHVC-PDASDMVHSEGLLIYKONIIPFMRVRKRVKVVTT 141
Db 36 YPFRICSIAGTDLMRDFDRDISCPYKSNAMSGFFIYKTNITETTFPVRYKNELT 95
Qy 142 STVYNGIYSDSITNQHTFYKSIETPETERKMDTIYQCFNSLRN--TGGNLLTYVDRDDIN 199
Db 96 QTSYRDVGTVYFLDRITVMGLAMPVYEANLNSRAQCYSAVAIKRPGDTVFSAYHEDNNKN 155
Qy 200 MTVPLOPVGTVDPVKRYGSOPELYLEPGFWGYSRRRTTVNCELMDMFARSNPPDFV 259
Db 156 ETLEFLPLNFKSVTKRFTITKEPYFARGPLW-LYSTSTSLNCIVTEATAKAKYPFSYFA 214
Qy 260 TATGDTVEMSPFWGSD-DHENKMHKPPWFSVNNYKV-DYQN--RGTVPLGKTRIFL 315
Db 215 LTTGEIVGSPFDGSGNGKFAEPLEK--LTILENTWIEDLMNGMGATTLVRKIAFL 271
Qy 316 DREYTLSEKHLKNNMNSYCPPLTWKAFYNGIQTSEHSGSYHFVANDITASFTSKEDMKEF 375
Db 272 EKGDTLFSWEIKENESVCLMKHWTTVTHGLRAETDTHYFISKELTAAFVASKESINLT 331
Qy 376 NTTYHCLNEEIKAEIKKY-AKVNSTHSGYGLKYFTDGGLYLVWQPLIQNRLLDANK 434
Db 332 DPQKTCIKNEFEKIITDVYMSDYNDAYSMMNGSYQIFKTTGDLILIQPLVQKSLM---V 387
Qy 435 LNNETYSRRSRQAEISTTDPMMETGNGAGGEYSSENSITVAQVOYAYDNLRIINNILE 494
Db 388 LEQGSVNLRRRDV-----DVKSRRHDLIVQLQYLYDLTKDYINDALG 431
Qy 495 DLKAWCREQHRAALVNNELSKINPTSMVMYNNRPVSAKRIGDVISVNCIVVDQTSVS 554
Db 432 NLAESWCLDQKRTITMLHELKSISSIVSEYVGRPISAQLHGDVLAISKIEVNQSSVQ 491
Qy 555 LHKSLRLLSA----SDEKCFSPRPVTFKFNWDS-TIYKQGLGVNNEILLTTTYLETQEN 609
Db 492 LYKSMRVVDAKGVRSCTMTCYNRPVTFSEFVNGTPEVVLGQLGLDNEILLGDRHTECEIP 551
Qy 610 TEYFQAKDMYIKNYEHLKTVPLSSITTLDTFIALNFTLLENVDFKVELYTRDEKRL 669
Db 552 STKIFLGNHAHVYDTHNNTSTIEDIEVLDAFIRLKDPLENADPKLHDLSPBELSR 611
Qy 670 SNVFDIETMFREYNVYAAQRVSGLRKDLLDLSTNRNQFVDAFGSLMDDLGAVGQTVNNAVS 729

Db 612 ANVFDLENILREYNSYK---SALYTTEAKIATNTPSVVNGINSFLOGLGAIGTGLGSVIS 668

Qy 730 GVATLFSSIVTGFINFIKPPFGGMLMIIVVIGLFAIYFLTKTKYETAPIKMIYP--- 786

Db 669 VTAGALGDIVGGVSVFLKNPFGGLMLILAIIVVVIIIVFVRQKHLVSLPIDMMFPYAT 728

Qy 787 -----EIDKLKEREKSEIAPISEEE 807

Db 729 NPVTTSVVTGTTVVKTPTSVKVDGTSVA-VSEKE 763

RESULT 13
T44186
probable glycoprotein B U39 [similarity] - human herpesvirus 6 (strain Z29)
C:Species: human herpesvirus 6
A:Variety: strain Z29
C>Date: 21-Jan-2000 #sequence_revision 21-Jan-2000 #text_change 05-May-2000
C:Accession: T44186
R:Dominguez, G.; Dambaugh, T.R.; Stamey, F.R.; Dewhurst, S.; Inoue, N.; Pellett, P.E.
J. Virol. 73, 8040-8052, 1999
A:Title: Human herpesvirus 6B genome sequence: coding content and comparison with human 1
A:Reference number: Z22734; MUID:99412318; PMID:10482553
A:Accession: T44186
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-830 <DOM>
A:Cross-references: EMBL:AF157706; PIDN:AAD49653.1
A:Experimental source: strain Z29; variant B
C:Genetics:
A:Note: U39
C:Superfamily: herpesvirus glycoprotein B

Query Match 22.9%; Score 1040.5; DB 2; Length 830;
Best Local Similarity 32.3%; Pred. No. 5.9e-54;
Matches 244; Conservative 155; Mismatches 298; Indels 59; Gaps 15;

Qy 83 FPYVCSASGVDFRQTDHVC-PDASDMVHSEGLLIYKONIIPFMRVRKRVKVVTT 141
Db 36 YPFRICSIAGTDLMRDFDRDISCPYKSNAMSGFFIYKTNITETTFPVRYKNELT 95
Qy 142 STVYNGIYSDSITNQHTFYKSIETPETERKMDTIYQCFNSLRN--TGGNLLTYVDRDDIN 199
Db 96 PYSYRDVGTVYFLDRITVMGLAMPVYEANLNSRAQCYSAVAIKRPGDTVFSAYHEDNNKN 155
Qy 200 MTVPLOPVGTVDPVKRYGSOPELYLEPGFWGYSRRRTTVNCELMDMFARSNPPDFV 259
Db 156 ETLEFLPLNFKSVTKRFTITKEPYFARGPLW-LYSTSTSLNCIVTEATAKAKYPFSYFA 214
Qy 260 TATGDTVEMSPFWGSD-DHENKMHKPPWFSVNNYKV-DYQN--RGTVPLGKTRIFL 315
Db 215 LTTGEIVGSPFDGSGNGKFAEPLEK--LTILENTWIEDLMNGMGATTLVRKIAFL 271
Qy 316 DREYTLSEKHLKNNMNSYCPPLTWKAFYNGIQTSEHSGSYHFVANDITASFTSKEDMKEF 375
Db 272 EKGDTLFSWEIKENESVCLMKHWTTVTHGLRAETDTHYFISKELTAAFVASKESINLT 331
Qy 376 NTTYHCLNEEIKAEIKKY-AKVNSTHSGYGLKYFTDGGLYLVWQPLIQNRLLDANK 434
Db 332 DPQKTCIKNEFEKIITDVYMSDYNDAYSMMNGSYQIFKTTGDLILIQPLVQKSLM---V 387
Qy 435 LNNETYSRRSRQAEISTTDPMMETGNGAGGEYSSENSITVAQVOYAYDNLRIINNILE 494
Db 388 LEQGSVNLRRRDV-----DVKSRRHDLIVQLQYLYDLTKDYINDALG 431
Qy 495 DLKAWCREQHRAALVNNELSKINPTSMVMYNNRPVSAKRIGDVISVNCIVVDQTSVS 554
Db 432 NLAESWCLDQKRTITMLHELKSISSIVSEYVGRPISAQLHGDVLAISKIEVNQSSVQ 491
Qy 555 LHKSLRLLSA----SDEKCFSPRPVTFKFNWDS-TIYKQGLGVNNEILLTTTYLETQEN 609
Db 492 LYKSMRVVDAKGVRSCTMTCYNRPVTFSEFVNGTPEVVLGQLGLDNEILLGDRHTECEIP 551
Qy 610 TEYFQAKDMYIKNYEHLKTVPLSSITTLDTFIALNFTLLENVDFKVELYTRDEKRL 669

```

Db      552 STKIFUSGNHAYVYDTYTHNSTPIEDIEVLDAPFIRLKDIPLENADFKLLDLYSPDELRS 611
QY      670 SNVFDLETFMREYNYAQRVSGLRKDLDDLSTNRNOFVDAFGLMDDLDGAVGQTVNVAVS 729
Db      612 ANVFDLENILREVNSYK---SALYTEAKIATNTPSYVNGINSFLOGLAGAIGTGLGSVIS 668
QY      730 GVATLFSSIVTGFINFKNPFGGMLMIIVVIGLFAIYFLTKTKIYETAPIKMIYP---- 786
Db      669 VTAGALGDIVGVVSLKPNPFGGGLMLILAIIVVVIIVFVRQKHVLSKPIDMWFYAT 728
QY      787 -----EIDKLKEREGKSEIAPISEEE 807
Db      729 NPVTTVSSVTGTTVTKPSVKDADGTSVA-VSEKE 763

RESULT 14
VGBE6S
N:Alternate names: glycoprotein B - human herpesvirus 6 (strain U102)
N:Contains: 58K glycoprotein; 64K glycoprotein
C:Species: human herpesvirus 6
A:Note: host Homo sapiens (man)
C:Date: 30-Sep-1992 #sequence_revision 19-Oct-1995 #text_change 16-Jul-1999
C:Accession: JQ1989; C40898
R:Ellinger, K.; Neipel, F.; Foa-Tomasi, L.; Campadelli-Piume, G.; Fleckenstein, B.
J. Gen. Virol. 74, 495-500, 1993
A:Title: The glycoprotein B homologue of human herpesvirus 6.
A:Reference number: PQ0616; MUID:93187613; PMID:9383182
A:Accession: JQ1989
A:Molecule type: DNA
A:Residues: 1-831 <ELL>
A:Cross-references: GB:Z18287
R:Teo, I.A.; Griffin, B.E.; Jones, M.D.
J. Virol. 65, 4670-4680, 1991
A:Title: Characterization of the DNA polymerase gene of human herpesvirus 6.
A:Reference number: A40898; MUID:91333007; PMID:1651403
A:Accession: C40898
A:Molecule type: DNA
A:Residues: 573-831 <TEO>
A:Cross-references: GB:M63804; NID:G325467; PIDN:AAA74632.1; PID:G325470
C:Superfamily: herpesvirus glycoprotein B
C:Keywords: glycoprotein; transmembrane protein
F:1-396/Product: 58K glycoprotein #status predicted <G58>
F:401-831/Product: 64K glycoprotein #status predicted <G64>
F:667-686/Domain: transmembrane #status predicted <TM1>
F:694-711/Domain: transmembrane #status predicted <TM2>
F:155,228,248,287,308,330,356,362,487/Binding site: carbohydrate (Asn) (covalent) #statu
Query Match 22.9%; Score 1039; DB 1; Length 831;
Best Local Similarity 32.1%; Pred. No. 7.2e-54;
Matches 244; Conservative 159; Mismatches 290; Indels 68; Gaps 16;

QY      83 FPRVCSAGVGDVFRFQTDHVC-PDASDMHSEGLLIYKNIIPFMRVRKRYKVVT 141
Db      36 YPRICSAKGTDLRFRDIDSCSPYKSNKMSGEFFIIYKTIETFPVRYTKKELTF 95
QY      142 STVYNGIYSDITNOHTFYKSTPEWETEKMDIYOCENSLRLN--TGGNLLTYVDRDIN 199
Db      96 QSSYRDVGUVYFLDRTVWGLAMPVTEANLVNSHAOCYSAVAMKRPDGTGVFSAFHEDNNKN 155
QY      200 MTFVLPQVDGVPDVKRYGSPQELYLEPGWFGSYRRRTTNVCELMDMFARSNPPDFDV 259
Db      156 NTLNLFPLNFKSITNKRFTITKEPIFARGPLW-LYSTSLNLCIVTEATAKAYFFSFA 214
QY      260 TATGDTVMSPTWSEGDHDKNH-----EKPWFVSVINNYKVY-DYQN--RGTVPLGK 310
Db      215 LTTGIVEGSPFFNG-----SNGKHFFAEPLK---LTILENTYMTIEDLMNGMGAYYLVR 267
QY      311 TRIFLDREYTLISWEKHLKNMSYCPILTLLKAFYNGIQTSEHSGSYHFVANDITASFTTSE 370
Db      268 KIAFLEKADTLFSWEIKEENESVCMKHWTTVTHGLRAETNETYHFISKELTAFAAPKE 327

```

```

QY      371 DMKEFTNTYHCLNEEIKAEIEKKY-AKVNSTHSHKYGLDKYFKTDGGLYLVMPQLIQNRLL 429
Db      328 SLNLTDPQKTCIKNEFEKIIINEVMSDYNDYVSMNGSYQIEKTTGDLLILIQPLVQKSLM 387
QY      430 DAKNKLNNETYSRRSRROAESTTDDPMEMTNGAGGEYSSENSITVAQVQYAYDNLRIRI 489
Db      388 FL-----EQGSEKIRRRRDV-----GDVKSRRHDLVYVQLQLYDFTLKGVI 427
QY      490 NNILEDLSKAWCREQHRRAALVNNELSKINPTSVMSMIYNRPVSAKRIGDVISVSNICIVVD 549
Db      428 NDALGNLAESWCLDQKTIITMLHELKSKSPSSIVSEVYGRFISAQLHGDVLAISKIEVN 487
QY      550 QTSVSLHKSRLLSA-----SDEKCFSPRPVTFKFMNDS-TIYKGLGVNNEILLTYYLE 604
Db      488 QSSVQLHKSMRVDAKGVRSCTMCYNRPLVTFSEVNSTPEVVPVPGQLGDLNELLGDHRTE 547
QY      605 TCQENTEYFQAKTDMYIKNYEHLKTVPLSSIITLDTFFIALNFTLLENVDFKIELYTR 664
Db      548 ECEIPSTKIFLUGNHAHVYTYTHNSTPIEDIEVLDAPFIRLKDIPLENADFKVLDLYSP 607
QY      665 DEKLSNVFDIETMFRYNYAQRVSGLRKDLDDLSTNRNOFVDAFGLMDDLDGAVGQTV 724
Db      608 DELSRANVFDLENILREVNSYK---SALYTEAKIATNTPSYVNGINSFLOGLAGAIGTGL 664
QY      725 VNAVSGVATLFPSSIVTGFINFKNPFGGMLMIIVVIGLFAIYFLTKTKIYETAPIKVI 784
Db      665 GSVISVTAGALGDIVGVVSLKPNPFGGGLMLILAIIVVVIIVFVRQKHVLSKPIDMM 724
QY      785 YP-----EIDKLKEREGKSEIAPISEEE 807
Db      725 FPYATNPVTTVSSVTGTTVTKPSVKDVGDTGTSVA-VSEKE 764

RESULT 15
VGBE6C
C:Species: murine cytomegalovirus (strain Smith [ATCC VR-194])
C:Date: 30-Jun-1992 #sequence_revision 30-Jun-1993 #text_change 09-Sep-1994
C:Accession: A41992; A40780; S23220
R:Rapp, M.; Messerle, M.; Buehler, B.; Tanneheimer, M.; Keil, G.M.; Koszinowski, U.H.
J. Virol. 66, 4399-4406, 1992
A:Title: Identification of the murine cytomegalovirus glycoprotein B gene and its expres
A:Reference number: A41992; MUID:92292266; PMID:1318410
A:Accession: A41992
A:Molecule type: mRNA
A:Residues: 1-928 <RAP>
A:Cross-references: GB:M86302
R:Elliot, R.; Clark, C.; Jaquish, D.; Spector, D.H.
Virology 185, 169-186, 1991
A:Title: Transcription analysis and sequence of the putative murine cytomegalovirus DNA
A:Reference number: A40780; MUID:92024072; PMID:1718083
A:Accession: A40780
A:Molecule type: DNA
A:Residues: 514-521, 'R', 523-526, 'E', 528-555, 'T', 557-561, 'V', 563-582, 'Q', 584-586, 'K', 588-
A:Cross-references: GB:M73549
C:Superfamily: herpesvirus glycoprotein B
C:Keywords: glycoprotein; transmembrane protein
F:1-36/Domain: signal sequence #status predicted <SIG>
F:37-928/Product: glycoprotein B #status predicted <GPB>
F:757-804/Domain: transmembrane #status predicted <TMN>
F:93,223,315,354,398,429,470,487,618,667/Binding site: carbohydrate (Asn) (covalent) #st
Query Match 22.7%; Score 1034; DB 1; Length 928;
Best Local Similarity 31.3%; Pred. No. 1.7e-53;
Matches 264; Conservative 155; Mismatches 324; Indels 100; Gaps 25;

QY      8 RGSVLALWLYQVAL-----YLSIAETGVTS-----PNTATWSTESPLT 48
Db      7 RGCSSSWYAMSTALATVINCILACTSEVIAAASPTGTTTAKAKTDTSETASATETATS 66
QY      49 GHYCTHDSHGERGNENRDS-----EQKNIIYGPSPTFFRYVCSAGVGDVF 97
Db      67 G-----AATGKKEATPTQASKITGTTIVPFVNEDMWSVDIDKYPYRVCMAVST-DLV 119

```

QY 98 RFQTDHYCPDASDMVH-SEGILLIYKONIIPFMFRVRYKRVVTTSTVYNGIYSDSITNQ 156
DB 120 RFGKSIDCINHPTKPTVOEGIMGVYKENIYAYFEVITYHK-DAIFORSYADTTNY 175
QY 157 H---TFYKSTIEP-WETEKMDTIYOCFNSL-RLNTGGNLLTYVDKDDINMTVFLOQVDPVGT 211
DB 176 FLGTSVTKMAFPWELDEVRNRCYSAASRLNGEVVYVAYHEDSYRNYTMWLVEDDYS 235
QY 212 PDVKRYSGQBELYLEPQWFGWSYRRRTTVCNCELMDFARSNPPDFVTTATGTVEMSPF 271
DB 236 KNSKRYVTTKSRVYHKGAWTW-RYTESCNMNCVVVTKARSNTPEFFVLSSGVEISPF 294
QY 272 WSGEDDENKMKHPFVSVINNYKVVDY-----QNRGTVPGLKTRIFLDBREYTLISWEK 326
DB 295 YNGENSEPFEEDTRNFWR--KNYTMKTYFGLAAPKKVPEL---MAFLEREDMTIGWEI 349
QY 327 HLKMSYCPGLTLWKAFYNGIOTHSYGFVANDITASFTTSKEDMKFNT-----T 378
DB 350 FPKQNVTCMDKKWGTVSRAIRTDINTGYHFSKGLTATFVASKRKI-DYNTTTEGKNYNT 408
QY 379 YHCLNEIKAEIKKYA-KVNSTHSKYGDLKYFTDGLLYVWQPLIONRLLDKKNLNN 437
DB 409 FRCVYDFEVRNRFVEYNEHTVKGDELMYRTTGGLIVLWQ-----LRAKSLHNL 462
QY 438 ETSYRRRRQAESTTDPM-MEMTNG-----AGGEYSSEN-----SITVAQVQYVDNLR 486
DB 463 EKFAALANNVSATASPPVTTAATENGTTVRSRRKRSFDNLDDVVVDTSYAOLOFTYDVLK 522
QY 487 IRINNILEDLSKAWCROHRAALVWNELSKINPTSVMSMYNRPVSAKRIGDVISVNCI 546
DB 523 DYINDALRNIMDWCROKRTAEMLKELSKINPNSIISAIRYERVTAKLAGDVIANSECV 582
QY 547 VVDOTSVLSHLKSLLSASDEK---CFSPRPVTFKFMNDSTIYKQGLGVANNEILLTYYL 603
DB 583 KVDQSSVKVLKDMRIF--QDGKVNCYSRPLVVFQFINSTKLESQGLGENNEIMLGTFT 640
QY 604 ETCQENTYFQAKTDMYIKYNYEHLKTVPLSSITTLDTFIALNFTLLENVDFKVIELYT 663
DB 641 ENCDTNSRKIFVVGTVGYEYRDRFRNVTSLIEDLVDTLIGLDIEPLENTDFKVLGLYS 700
QY 664 RDEKRLNVPDIEPTMFREYNVYQVSGRLKDLDLSTNRNQFVDFAGSLMDLGLAVGQT 723
DB 701 KGEIRASNVSLDIMEYNSQKHIRTLSAKVND---NTPSYLLGLDTFMQGLGVAGKG 757
QY 724 VNAVSGVATLFFSSIVTGFINFINKPEGMLMIIVIGVLFALYFLTKTKIYETAPIKM 783
DB 758 IGVAIGAVGAVSVNAVTVGFLTNPFGGFTTILLVIGLAVVYLIF-----TRDVEY 810
QY 784 IYP 786
DB 811 FFP 813

RESULT 16

D48349 glycoprotein B precursor - saimirine herpesvirus 1 (strain MV-5-4-PSL)

C;Species: saimirine herpesvirus 1

C;Date: 17-Feb-1994 #sequence_revision 17-Feb-1994 #text_change 09-Sep-1994

C;Accession: D48349

R;Eberle, R.; Black, D.

Arch. Virol. 129, 167-182, 1993

A;Title: Sequence analysis of herpes simplex virus gB gene homologs of two platyrrhine m

A;Reference number: A48349; MUID:93228440; PMID:8385913

A;Molecule type: DNA

A;Residues: 1-920 <EBE>

A;Note: sequence extracted from NCBI backbone (NCBIN:129066, NCBI:129068)

C;Superfamily: herpesvirus glycoprotein B

C;Keywords: Glycoprotein; transmembrane protein

F;1-35/Domain: signal sequence #status predicted <SIG>

F;36-920/Product: glycoprotein B #status predicted <MAT>

F;576-592/Region: hydrophobic

F;761-780/Domain: transmembrane #status predicted <TM1>

F;787-804/Domain: transmembrane #status predicted <TM2>

F;98,119,152,409,441,683,733/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 19.7%; Score 894; DB 1; Length 920;

Best Local Similarity 28.9%; Pred. NO. 3.4e-45;

Matches 242; Conservative 155; Mismatches 359; Indels 82; Gaps 21;

QY 60 ERGNENRUSEEQ-----NKNIYGPSPTFFRYVCSASGVGDVFPQTDHYCPD-ASDMVHS 114

DB 95 QRANKTAEDARAOLROSVRQIRAEANATSMFYQPPPTGATVVOQFEPFPCPDVAAQKNFT 154

QY 115 EGTLLIYKONIIPFMFRVRYKRVVTTSTVY-----NGIYSDSITNQHITFKSTEPW 166

DB 155 EGIYVIFKENIADPKYKATMYKEIIVTQWQSGRVLQTLGLND-----RAPVPF 205

QY 167 E--TERKMDTIYOCFNSLRLNTGNLLTYVDKDDINMTVFLOQVDPVTKRYGSOPELY 224

DB 206 EETDVINAKGLCRSDVTYVRSQRRVTAYDRDEWGREVKLVPSKSTSPNSRGWVTTDRMY 265

QY 225 LEPCWFGSYRRRTTVCNCELMDFARSNPPDFVTTATGTVEMSPFWS-GEDDH-ENKM 282

DB 266 -AFNAHAGFYKAGTTVCIVEEVEARSAYPSNFVLATGDFVYVSPFYGLGEDAHRAYNA 324

QY 283 HEKPFVSVINNYKVVDYQNRGTVPGLKTRIFLDBREYTLISWEKHLKMSYCYPLTLWKAF 342

DB 325 YSADRKFQV-DGFFPRDLDSGETAPEVVRNLLTTPKFTIGWDWKPKDPSVCVTKWEV 383

QY 343 YNGIOTHSYGFVANDITASFTTSKEDMKFNTTYH-----NVTQYPPQRIELSDCVAREAAQVDAI 435

DB 384 EEMWRAEYSGTFRFTSSLSATFT-----NVTOYPPQRIELSDCVAREAAQVDAI 435

QY 394 YA-KVNSTHSKYGDLKYFTDGLLYVWQPLIONRLLDKKNLNNETYSR-----RERQ 448

DB 436 YARRYNASHVKVGLQYLAQGLVYVQFLISNSLA-----EMYLEABARALEPA 487

QY 449 ESTTDPMEMTGNAGGEYSSENSITVAOVAYDNLRIINNILEDLSKAWCREQHRAA 508

DB 488 PLPTTAPAEA--GSRGTLSSTQSFARLQFTYDHIOKHVNEMLGRIAAAWCOLNQEL 545

QY 509 LVWNELSKINPTSVMSMYNRPVSARKRIGDVISVNCI VVDOTSVLSHLKSLLSASDEK 568

DB 546 VLMNEARKLNPAIASATVRRVVGARMLGDVMAVSTCIAPVPHNVIMQNSMR-LPARPKT 604

QY 569 CFSRPVTFKFMNDSTIYKQGLGVANNEILLTYYL ETCQENTYFQAKTDMYIY-KNYE 627

DB 605 CYARPLVSFRYADEGLIEGOLGEDNEIRLEQNLEPCTVGHKRYF-VFGDGVYFEEYA 663

QY 628 HLKTVPLSSITTLDTFIALNFTLLENVDFKVIELYTRDEKRLSNVDFDIETMFREYNVY 687

DB 664 YSHQVSRADVVPVSTFVDNLNLTWLEHBFLOVYTRQEI KDSGLLDYAEVQRNQWAL 723

QY 688 RVSGLRKDLDLSTNRNQFVDFAGSLMDLGLAVGQTVNAVSGVATLFFSIVTGFINF 747

DB 724 RFSDI--DHIINDTTNAALMDGLFRFFDGLGAAGQAIGRAVLGVTAEAVISVSGVSFLS 781

QY 748 NPGGMLMIIVIGVLFALYFLTKTKIYETAPIKMIY-----EIDKLKERE 795

DB 782 NPFAGALVGLLVLGATTAFFALRYIMRLANRPMRALYPTTHGICAKAKASASGEP 841

QY 796 GKSEIAPISEEELERIVLAWHIHQNSHMETKTRKDPKDSILTRAQ-----MLRKRS 848

DB 842 GPGGIEDFEAKLEAEARTWIKYMTLVSAHERTAHAKKRGTSARISHLTDMVLRKN 899

RESULT 17

VGBERB

glycoprotein B precursor - Marek's disease virus (strain RB1B)

C;Species: Marek's disease virus

C;Date: 31-Dec-1991 #sequence_revision 31-Dec-1991 #text_change 16-Jun-2000

C;Accession: A32402; B32402

R;Ross, L.J.N.; Sanderson, M.; Scott, S.D.; Binns, M.M.; Doel, T.; Milne, B.

J. Gen. Virol. 70, 1789-1804, 1989

QY	732	AI	FE	SS	VI	GF	IN	FK	PF	GG	ML	II	VI	GV	FA	IF	FK	TI	YE	TA	PI	KM	YI	PE	DK	791
		:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	
		:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	
DB	714	AG	AI	VS	TI	GS	VA	FM	SP	FG	AL	GI	II	AG	LV	AA	FL	AY	VN	KL	KS	NM	KAL	YPM	TV	773

QY 528 NRPVSAKRIQGVILSVSNICIVDDQTSVSHKSKRLTSLASDEKCFSPRPVTFKFMNDUSTIYK 58
Db 594 ERRVSARLLGLDAVAVTQCNVSSGHVYIQNSMR-VTGSSITCYSRPLSVSRALNDSEYIE 652


```
Db      792  FAFYVMRLQSNPKALYPLTTKELKNPTNPDSAGEGEGDFDEAKLAAREMIRYMAL 851
Qy      815  MH1HQONSHMETKTRKDPKDSILT-RAQNML-----RKRGSGYNLKNAE 857
Db      852  VSAMERTEH---KAKKGTGSRLLSAKVTDVMVRKRRTNTYQVFNKD 895

RESULT 22
VGBEW7
glycoprotein B precursor - human herpesvirus 1 (strain 17)
C:Species: human herpesvirus 1
C>Date: 31-Dec-1989 #sequence_revision 31-Dec-1989 #text_change 16-Jun-2000
C:Accession: I30084
R:McGeoch, D.J.; Dalrymple, M.A.; Davison, A.J.; Dolan, A.; Frame, M.C.; McNab, D.; Perz
J. Gen. Virol. 69, 1531-1574, 1988
A>Title: The complete DNA sequence of the long unique region in the genome of herpes sim
A:Reference number: A30083; MUID:88274327; PMID:2839594
A:Accession: I30084
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-904 <MC>
A:Cross-references: GB:X14112; GB:D00317; GB:D00374; GB:S40593; NID:g1944536; PIDN:CAA32
C:Genetics:
A:Gene: UL27
C:Superfamily: herpesvirus glycoprotein B
C:Keywords: glycoprotein; transmembrane protein
F:1-30/Domain: signal sequence #status predicted <SIG>
F:31-904/Product: glycoprotein B #status predicted <MAT>
F:727-746/Domain: transmembrane #status predicted <TM1>
F:752-771/Domain: transmembrane #status predicted <TM2>
F:774-795/Domain: transmembrane #status predicted <TM3>
F:87,141,255,398,430,478,489,674,819,888/Binding site: carbohydrate (Asn) (covalent) #82
F:116-573,133-529,207-271,364-412,596-633/Disulfide bonds: #status predicted

Query Match      18.5%; Score 841; DB 1; Length 904;
Best Local Similarity 27.7%; Pred. No. 4.7e-42;
Matches 242; Conservative 179; Mismatches 345; Indels 108; Gaps 28;

Qy      35  PPNATWSTESPLTGHVYTHDSSHGE-RGNENNRDSEONKNYGSPTPYRVCSASGV 93
Db      80  PPRPA-----GDNATVAAGHATLRHLRDLKAENTDANFY-----VCPPTG 121
Qy      94  GDVRFQTDHVCDDASD-MVHSEGLIYKQNIIPFMRVRKVKVTTSTV----- 144
Db      122  ATVVQFQPRRCPRPGQNYTEGIAVFNKENTAPIKFKATMYKQVTVSQVWFHGRYSQ 181
Qy      145  YNGIYSDITNQHTFYKSIIEPWE--TEKMDTIYQCFNSRLNTGGNLLTYVDRDDINMTV 202
Db      182  FMGIFED-----RAPVPFEEVIDKINAKGVCSTAKYVNNLETTAFHRDDHETDM 232
Qy      203  FLOPVDGV-----TPDVKRYGSDPELYLEPGWFWGVSRYRRRTVNCELMDMFARSNPP 254
Db      233  ELKPANAATRTSRGWHTTDLKYNRSVEAF-----HRYGTTVNCIVEEVDARSVYP 283
Qy      255  FDFPVTATGTVENSPWS-GEEDH-ENKMKHEKPFVSVINNY-KVVDYQNRGTVP LGKT 311
Db      284  YDEEVLATGDFVNSPFGYREGSHTHTSYAADRFXQVDFGFIARDLTTKARATAP--TT 341
Qy      312  RIFLDREYILTSWEKHLKNMNSCYPLTLWKAFYNGIQTEHSGSYHFVANDITASFTT--- 367
Db      342  RNLLTTPKFTVANDVPKRSVCTMTKQVDEMLRSEYGGSPFSSDAISTFTTNLTE 401
Qy      368  ---SKEDMKFNTYHCLNEEIKAEIEKKYA-KVNSTHSKYGDLKYFKTGGGLYLVNQPL 423
Db      402  YPLSRVDLGD-----CIGKDARDAMDRIFARRYNATHIKVGQPOYYLANGGELIAQPL 455
Qy      424  IQNELLDKKNLNNETYSRRSRQAESTTDPMMEMTCNGAGGEYSSENSITVAQVOYAD 483
Db      456  LSNLTAE-----LYVREHLRQSRKPPPTPPPGASANASVERIKTSSIEFARLQPTYN 511
Qy      484  NLRIRINILEDLSKAWCREQHRALVNMELSKINPTSVMSIYNRPVSAKRIGDVISVS 543
```

```
Db      512  HIQRHVNMDLGRVAIAJAWCEQLNHLELTWNEARKLNPNALIASATVGRVRVSRMLGDVMAVS 571
Qy      544  NCIVVDQTSVLSHLKSLRLLSASDEKCFSPPTVTFKEMNDSTIYKGGOLGVNNEILLTTTTL 603
Db      572  TCVFVAADNVIVQNSMR-ISSRPGACYSRPLVSFRYEDOGQPLVEGOLGNNELRLRDAI 630
Qy      604  ETCQENTETYYFOAKTDMYIYKNVEHLKTVPLSSITTLDTFFIALNFTLLENVDFPKVIELYT 663
Db      631  EPTCVGHRRYFTFGGYYVFEVAYVSHQSLRADITTVSTVFIDNLITMLEDHEFPVLEVYT 690
Qy      664  RDEKRLSNVFDIETMPREYNYYAQRVSGRLKOLLDLSTNRNQF--VDAFGLMDDLG-AV 720
Db      691  RHEIKDGLLDYTEVQRRNQLDLRPAI-DTVIHADANAAMPAGLGAFFEGMGDLGRAV 749
Qy      721  GQTVNVNAGSVATLPSSIVTGFNFKNPFGGMLMIIIVIGLVFAIYFLTKTKIVETAP 780
Db      750  GKVMVGIVGWV-----SAVSGVSSPNSNPFGLAVGLLVLAGLAAFAFFAFRYVMRQSNP 805
Qy      781  IKMIYPEIDK-LK-----EREGKSEIAPISE-BELERIVLAMHIHQONSHMETK 827
Db      806  MKALYPLTTKELKNPTNPDSAGEGEGDFDEAKLAAREMIRYMALVSAMERTEH--K 862
Qy      828  TRKDPKDSILT-RAQNML---RKRGSGYNLKNAE 857
Db      863  AKKKGTSALLSAKVTDVMVRKRRTNTYQVFNKD 896

RESULT 23
VGBEB2
glycoprotein B precursor - human herpesvirus 2 (strain 333)
C:Species: human herpesvirus 2
C>Date: 31-Mar-1989 #sequence_revision 31-Mar-1989 #text_change 16-Jul-1999
C:Accession: A26790; A45928
R:Stuve, L.L.; Brown-Shimer, S.; Pachl, C.; Najarian, R.; Dina, D.; Burke, R.L.
J. Virol. 61, 326-335, 1987
A>Title: Structure and expression of the herpes simplex virus type 2 glycoprotein gB gene
A:Reference number: A26790; MUID:87112925; PMID:3027364
A:Accession: A26790
A:Molecule type: DNA
A:Residues: 1-904 <STU>
A:Cross-references: GB:M15118; NID:G330256; PIDN:AAA45837.1; PID:G330257
R:Zwaagstra, J.C.; Leung, W.C.
Can. J. Microbiol. 33, 879-887, 1987
A>Title: The nucleotide sequence of herpes simplex virus type 2 (333) glycoprotein gB2 an
A:Reference number: A45928; MUID:88079667; PMID:2446730
A:Accession: A45928
A:Molecule type: DNA
A:Residues: 1-34, 'AWPTV', 42-307, 'T', 309-481, 'R', 483-609, 'M', 611-664, 'R', 666-904 <ZWA>
A:Cross-references: GB:M24771; NID:G341245; PIDN:AAA60540.1; PID:G623400
A>Note: the authors translated the codon ATG for residue 610 as Ile
C:Superfamily: herpesvirus glycoprotein B
C:Keywords: glycoprotein; transmembrane protein
F:1-22/Domain: signal sequence #status predicted <SIG>
F:23-904/Product: glycoprotein B #status predicted <MAT>
F:724-744/Domain: transmembrane #status predicted <TM1>
F:749-769/Domain: transmembrane #status predicted <TM2>
F:772-792/Domain: transmembrane #status predicted <TM3>
F:82,136,250,393,425,473,486,671/Binding site: carbohydrate (Asn) (covalent) #status predi
F:111-570,128-526,202-266,359-407,593-630/Disulfide bonds: #status predicted

Query Match      18.5%; Score 839; DB 1; Length 904;
Best Local Similarity 28.3%; Pred. No. 6.2e-42;
Matches 246; Conservative 164; Mismatches 367; Indels 92; Gaps 27;

Qy      33  TSPNTATWSTESPLTGHVYTHDSSHGERGNENNRDSEONKNYGSPTPYRVCSASG 92
Db      76  TPEPD---ANATVAAGHATLR-AHLREIKVENADAQ-----EYVCPPT 115
Qy      93  VGVDFEFTDHCVPDASD-MVHSEGLIYKQNIIPFMRVRKVKVTTSTV----- 144
Db      116  GATVVOFEQPRRCPRPGQNYTEGIAVFNKENTAPIKFKATMYKQVTVSQVWFHGRYS 175
Qy      145  YNGIYSDITNQHTFYKSIIEPWE--TEKMDTIYQCFNSRLNTGGNLLTYVDRDDINMT 201
```

```
Db 176 QFMGIFED-----RAPVPEEVIDKINAGVCKSTAKYVRNNMETTAFHRDDHETD 226
Qy 202 VELQPDVGTTPDKRYGSOPELYLEPGFWGSRRTTUNCCELMDFARNPPFFDFVTA 261
Db 227 MELKPAK-VATRTSRGWHHTDLKYNPSRVEAFHRYGTTVNCIVEEDARSVPYDFEFLA 285
Qy 262 TGDVTMSPFWS-GEDDH-ENKMHKPMFVSVINNY-KVVDYQNRGTVPGLKTRIFLDRE 318
Db 286 TGFVYMSPPGYREGSHTEHTSYAADRFKQVDFYARDLTTKARATSP--TTRNLLTTP 343
Qy 319 EYTLSEKHLKMSYCPFLWKAFYNGIQTEHSGSYHFVANDITASFTT-----SKED 371
Db 344 KFTVAMDWPVKRPAPVCTMTKQVEDEMLRAEYGGSPRFSDDAISITFTTNLTQYLSRVD 403
Qy 372 MKEFNNTYHCLNEEIKAEIEKKYA-KVNSTHSGYGLKYFKTGDGLYLWQPLIONRLLD 430
Db 404 LGD-----CIGRDAREALDRMFARKYNATHIKVGOPOQYILATGGLFLAYOPLSNLAE 457
Qy 431 AKNKLNNETYSR-RSRROAESTTDPMMEM-TGNAGAGEYSSSENSITVAQVQYAYDNLR 488
Db 458 ----LVREYMRQDRKPRNATPAPLREAPSANASVERIKTSSIEFARLQFTYNHIQRH 513
Qy 489 INNILEDKAMCREQHRALVWNLKINPTSVMSIYNRPVSAKRIGDVISVNCIYV 548
Db 514 VNDMLGRIAVAWCELQNHLETLWNEARKLNPNAIASATVGRRVSRARMLGDVMAVSTCPV 573
Qy 549 DQTSVSLHSLRLLSASDEKCFSPRPVTFKPMNDSTIYKQGLGVNNEILLTLYLETCOE 608
Db 574 APDNVIVQNSMR-VSSRPGTCYSRPLVSFRYEDQGPLIEGOLGENNELLTRDALEPCTV 632
Qy 609 NTEYFFQAKTDMYIYKNYEHKLTVPPLSSITLDTFIALNFTLLENVDFKVIELTRDEKR 668
Db 633 GHRGVYFEGGYVVEYAYSHQLSRADVTTVSTFIDLNTMLEDHFPVLEVYTRHEIK 692
Qy 669 LSNVFDIETMFREYNYAQRVSGRLKDLDLSTNRNQFVD--AFGSLMDDLG-AVGQTVV 725
Db 693 DSGLLDYTEVQRNQLHDLRFADI-DTVIRADANAAMFAGLCAFFEGMGDLGRAVGKVM 751
Qy 726 NAVSGVATLFPSSITVTFINFKNPPFGMLMIIVIGVLFALYFLTKTKIYETAPIKMIY 785
Db 752 GVVGGVV----SAVGSVSSFMSPFGALAVGLLVLAGLVAAFFAFRVLQLOQNPMAKY 807
Qy 786 PEIDKLKEREKSEAPISEB-----ELERIVLAMHIHQONSHMETTKRDP 832
Db 808 PLTTKELKTSDPGCVGGEGBEGGDEAKLAAREMIRYALVAMSAMERTEHKARKKG 867
Qy 833 KDSIL-TRAQNM-LRKR--SGYSNLKNAE 857
Db 868 TSALLSSKVTNMVLKRKRKARYSPLHNE 896
```

RESULT 24

VBGEK2

glycoprotein B precursor - human herpesvirus 2 (strain HG52)
C.Species: human herpesvirus 2
C.Date: 31-Mar-1988 #sequence_revision 31-Mar-1988 #text_change 16-Jul-1999
C.Accession: A25611
R:Bzik, D.J.; Debroy, C.; Fox, B.A.; Pederson, N.E.; Person, S.
Virology 155, 322-333, 1986
A.Title: The nucleotide sequence of the gB glycoprotein gene of HSV-2 and comparison with
A.Reference number: A25611, MUID:87071654; PMID:3024391
A.Accession: A25611
A.Molecule type: DNA
A.Residues: 1-904 <BZI>
A.Cross-references: GB:M14923; NID:G330254; PIDN:AAA66440.1; PID:G330255
R.Norais, N.; Tang, D.; Kaur, S.; Chamberlain, S.H.; Maslars, F.R.; Burke, R.L.; Marcus,
J. Virology 70, 7379-7387, 1996
A.Title: Disulfide bonds of herpes simplex virus type 2 glycoprotein gB.
A.Reference number: A58366; MUID:97048015; PMID:8892856
A.Contents: annotation; tryptic peptide disulfide bond assignments
C.Superfamily: herpesvirus glycoprotein B
C.Keywords: glycoprotein; transmembrane protein

F:1-22/Domain: signal sequence #status predicted <SIG>
F:23-904/Product: glycoprotein B #status predicted <MAN>
F:724-744/Domain: transmembrane #status predicted <TM1>
F:749-769/Domain: transmembrane #status predicted <TM2>
F:772-792/Domain: transmembrane #status predicted <TM3>
F:82,136,250,393,425,473,486,671/Binding site: carboxylate (Asn) (covalent) #status pr
F:111-570,128-526,202-266,359-407,593-630/Disulfide bonds: #status experimental

Query Match 18.4%; Score 837; DB 1; Length 904;
Best Local Similarity 28.1%; Pred. No. 8.1e-42;
Matches 244; Conservative 167; Mismatches 366; Indels 92; Gaps 27;

```
Qy 33 TSPNTATWSTESPLTGHYTHDSSHGGRGNENRDEEONKNIYGSPTPYRVCSASG 92
Db 76 TTPPD-----ANATVAAGHATVR--AHLREIKVENADAQ-----FYVCPPT 115
Qy 93 GVDVFRFOTDHYVCPDASD-MVHSGGILLIYKONIIPMFRVRKYRKVVTTTV----- 144
Db 116 GATVVQFEQPRRCPTRPREGQNYTEGIAVVFKEKNIAPYKATMYKDVTVSQVWFGHRS 175
Qy 145 -YNGIYSDSIITNQHFTFYKSIETPWE--TEKMDTIYQCENSLRNLNTGGNLLTYVDRDDINMT 201
Db 176 QFMGIFED-----RAPVPEEVIDKINAGVCKSTAKYVRNNMETTAFHRDDHETD 226
Qy 202 VELQPDVGTTPDKRYGSOPELYLEPGFWGSRRTTUNCCELMDFARNPPFFDFVTA 261
Db 227 MELKPAK-VATRTSRGWHHTDLKYNPSRVEAFHRYGTTVNCIVEEDARSVPYDFEFLA 285
Qy 262 TGDVTMSPFWS-GEDDH-ENKMHKPMFVSVINNY-KVVDYQNRGTVPGLKTRIFLDRE 318
Db 286 TGFVYMSPPGYREGSHTEHTTYAADRFKQVDFYARDLTTKARATSP--TTRNLLTTP 343
Qy 319 EYTLSEKHLKMSYCPFLWKAFYNGIQTEHSGSYHFVANDITASFTT-----SKED 371
Db 344 KFTVAMDWPVKRPAPVCTMTKQVEDEMLRAEYGGSPRFSDDAISITFTTNLTQYLSRVD 403
Qy 372 MKEFNNTYHCLNEEIKAEIEKKYA-KVNSTHSGYGLKYFKTGDGLYLWQPLIONRLLD 430
Db 404 LGD-----CIGRDAREALDRMFARKYNATHIKVGOPOQYILATGGLFLAYOPLSNLAE 457
Qy 431 AKNKLNNETYSR-RSRROAESTTDPMMEM-TGNAGAGEYSSSENSITVAQVQYAYDNLR 488
Db 458 ----LVREYMRQDRKPRNATPAPLREAPSANASVERIKTSSIEFARLQFTYNHIQRH 513
Qy 489 INNILEDKAMCREQHRALVWNLKINPTSVMSIYNRPVSAKRIGDVISVNCIYV 548
Db 514 VNDMLGRIAVAWCELQNHLETLWNEARKLNPNAIASATVGRRVSRARMLGDVMAVSTCPV 573
Qy 549 DQTSVSLHSLRLLSASDEKCFSPRPVTFKPMNDSTIYKQGLGVNNEILLTLYLETCOE 608
Db 574 APDNVIVQNSMR-VSSRPGTCYSRPLVSFRYEDQGPLIEGOLGENNELLTRDALEPCTV 632
Qy 609 NTEYFFQAKTDMYIYKNYEHKLTVPPLSSITLDTFIALNFTLLENVDFKVIELTRDEKR 668
Db 633 GHRGVYFEGGYVVEYAYSHQLSRADVTTVSTFIDLNTMLEDHFPVLEVYTRHEIK 692
Qy 669 LSNVFDIETMFREYNYAQRVSGRLKDLDLSTNRNQFVD--AFGSLMDDLG-AVGQTVV 725
Db 693 DSGLLDYTEVQRNQLHDLRFADI-DTVIRADANAAMFAGLCAFFEGMGDLGRAVGKVM 751
Qy 726 NAVSGVATLFPSSITVTFINFKNPPFGMLMIIVIGVLFALYFLTKTKIYETAPIKMIY 785
Db 752 GVVGGVV----SAVGSVSSFMSPFGALAVGLLVLAGLVAAFFAFRVLQLOQNPMAKY 807
Qy 786 PEIDKLKEREKSEAPISEB-----ELERIVLAMHIHQONSHMETTKRDP 832
Db 808 PLTTKELKTSDPGCVGGEGBEGGDEAKLAAREMIRYALVAMSAMERTEHKARKKG 867
Qy 833 KDSIL-TRAQNM-LRKR--SGYSNLKNAE 857
Db 868 TSALLSSKVTNMVLKRKRKARYSPLHNE 896
```



```
QY 418 LVWQPLI-----QNRLLDAKNKL-----NNETYGRRRRRQAESTTDDMMMENT 459
Db 475 IAFRPMLSNELARLYLNLVSRNRTYDLKNLLNPNANNNTTTRRRSLLSVPEQPTQD 534
QY 460 G-----NGAGGEYSSE-----NSITVAQVQAYADNLRIRINNI 492
Db 535 GVHRGQILHRLHRAVEATAGTDSNVNTAKQLELKTSTSSIEFAMQLQFAYDHIQSHVNM 594
QY 493 LEDLSKAMCREQHRALVNNELSKINPTSVMSMIYNRPVSARKIGDVISVSNVICVVDDOTS 552
Db 595 LSRIATAWCTLQNKERTLWNEVKINPSAIVSATLDERVAARVLDGVDAITHCAKI-EGN 653
QY 553 VSLHSLRLLSASDEKCFSPVPVF---KFMNDSTIYKQGLGVNNEILLTTTLYLETCOEN 609
Db 654 VYLQNSMR--SMDNSNTCYSRPVPVTFITKNANNRGSIEGOLGEENEIIFTERKLIPECALN 711
QY 610 TEYFQAKTDMYIKNYEHLKVPPLSSITLTFTFALNFTLLENVDFKVIELYTRDEKRL 669
Db 712 QKRYFKGKEVYVYENYTFVRKVPPEITEVISTYVELNLTLLEDRFLPLEVYTRAELED 771
QY 670 SNVFDIETMFREYNYAORVSGRLKDLDDLSTNRNQFVDAFGLSMDLGAVGOTVUNAVS 729
Db 772 TGLLDYSEIQRRNQLHALRFYDI--DSVVNVNDTAVIMQGIASFPGKLGKVGAEVGTIVL 829
QY 730 GVATLFSSIVTGFINFKNPFPGMLMIIVIGVLFALYFLTKTKIYETAPIKMIYPEID 789
Db 830 GAAGAVSVTVSGIASFLNPNFGLAIGLLVIAGLVAAPFAFYRVQMIRSNPMKALYPITT 889
QY 790 KLKEREGKSEIAPISE-----BELERIVLAMTHQONSHMETKTRDKPDS--- 835
Db 890 KALKNKATSYGQNEEDGSDGDFDEAKLEAREMIKYMVSALERQEKKAIK--KNSGVG 947
QY 836 -ILTRAQNMLRKRSQ--YSNLKNAESVE 860
Db 948 LIASNVSKLALRRGPKYTRLQONTME 975

RESULT 27
VGBESA
Glycoprotein B precursor - simian herpesvirus SA8 (strain B264)
C:Species: simian herpesvirus SA8
C:Date: 30-Sep-1992 #sequence_revision 30-Sep-1992 #text_change 16-Jul-1999
C:Accession: JQ1332
R:Bochers, K.; Weigelt, W.; Buhk, H.J.; Ludwig, H.; Mankertz, J.
J. Gen. Virol. 72, 2299-2304, 1991
A:Title: Conserved domains of glycoprotein B (gB) of the monkey virus, simian agent 8, i
A:Reference number: JQ1332; MUID:91374035; PMID:1895066
A:Accession: JQ1332
A:Molecule type: DNA
A:Residues: 1-885 <BOR>
A:Cross-references: EMBL:X56935; NID:G60438; PIDN:CAA40256.1; PID:G60439
C:Superfamily: herpesvirus glycoprotein B
C:Keywords: glycoprotein; transmembrane protein
F:1-28/Domain: signal sequence #status predicted <SIG>
F:129-885/Product: glycoprotein B #status predicted <SMAT>
F:712-732/Domain: transmembrane #status predicted <TM1>
F:737-752/Domain: transmembrane #status predicted <TM2>
F:760-780/Domain: transmembrane #status predicted <TM3>
F:68,122,379,411,659/Binding site: carbohydrate (Asn) (covalent) #status predicted
F:197-558,114-514,188-252,345-393,581-618/Disulfide bonds: #status predicted

Query Match
Best Local Similarity 27.1%; Pred. No. 1.4e-41;
Matches 253; Conservative 164; Mismatches 392; Indels 124; Gaps 26;

QY 5 LKLRG---SVLAWLYQVALYSLSIAETGVTSPNTATWSTESPLTGHYTHDSHGER 61
Db 1 MRPRGTPPSFLPLVLLALV---IAAGRAAPAAAAAPTADPAATPALPEDEBVPD 56
QY 62 GNNENRDEEQNKNIYGSPTF-----PVRVCSASGVGVDFRQTDHVCPD 107
Db 57 GEGVATPAPAANASVEAGRATLREIDREIKARDGDATFYVCPPTGTATVQVQEPQPCPR 116
```

```
QY 108 ASD-MVHSEGILLIYKQNIIPFMFRVRKYRKVVTSTV-----YNGIYSDSITNOHT 158
Db 117 APDQNYTEGIAVFKENIAPYKFKATWYKDVTVSQWFGHRYSQFNGIED----- 169
QY 159 FYKSIEPWETEKMDTIYO---CFNSLRLLNTGNNLLTYVDRDINNMTVPLOPVGVTDPVK 215
Db 170 --RAPVPPE-EVMOKINAKGVCRSTAKVYRNMMESTAFHRDDHESDMALKPAKAA-RTS 225
QY 216 RYGSQPELYLPFGWFGSYRRRTTNCCLMDMFARSNPDPFVFTATGDTVEMSPFWSGE 275
Db 226 RGWHTTDLUKYNPARVEAPHRYGTTVNCIIVEEARSVPYDFEVLATGDFVMSPFYGR 285
QY 276 D-DH-ENKMHKPPFVSVINNY-KYVDYQNRGTVPGLKTRIFLDREYETLSMEKLNKMS 332
Db 286 DSGHEHTAYAADRFQVDGYVERDLSGRRAAAPV--TRNLLTTPKTVGMDWAPKPS 343
QY 333 YCPLTLWKAIFYNGIQTEHSGSYHFVANDITASFTT-----SKEDMKEFNTYHCLNEE 385
Db 344 VCTLTKREVDEMRAEYGPSPFRSSAALSTTFTANRTEYALSRVDLAD-----CVGRE 397
QY 386 IKAETEKY-AKVNSTHSKYGLKYEKTDGGLYLVMOPLIONRLLDKNKLNNETYSRRS 444
Db 398 AREAADIFLFRYNGTHVKVQVYLYATGFLIAYQPLLSNALV-----ELYVREL 449
QY 445 RROAESTTDPMMEMTNGAGGE-----YSSENSITVAQVQAYADNLR 486
Db 450 VR-----EQTRRPAGDPCGEAATPGPSVDPSPVERIKTTSSEFARLQFTYDHIQ 499
QY 487 IRINNILEDLSKAWCRQHRALVNNELSKINPTSVMSMIYNRPVSARKIGDVISVNCI 546
Db 500 RHVNDMLGRIATAWCELQNRELTLWNEARRLNPGAISATVGRRVSRMGLDVMVAVSTCV 559
QY 547 VVDQTSVSLHKSRLLSASDEKCFSPVPVTFKPMNDSTIYKQGLGVNNEILLTTTLYETC 606
Db 560 PVAPDNVIMQNSIG-VAARPQTCYSRPLVSPRYEADGGLVEGJEDNEILERDALPCC 618
QY 607 QENTEYFQAKTDMYIKNYEHLKTVPLSSITLTFTFALNFTLLENVDFKVIELYTRDE 666
Db 619 TVGHRVYTFPCAGYVYFEEYAYSHQLGRAVDVTVSTVFNLNLTMLEDEHFVPLEVYTRQE 678
QY 667 KRLSNVFDIETMFREYNYAORVSGRLKDLDDLSTNRNQFVDAFGSLMDDI.GAVGQTVVN 726
Db 679 IKDSGLLDYTEVQRRNQLHALRFADI-DTVIKADAAHALFAGLY-SFPEGLGDDVGRAVGK 736
QY 727 AVSGVATLFSIVTGFINFKNPFPGMLMIIVIGVLFALYFLTKTKIYETAPIKMIYP 786
Db 737 VVMGIVGWSAVSGVSFLSNPFGALAVGLLVLGAAAAFFAFPRYVNRQLQNPMLKALYP 796
QY 787 EIDKLKEREGKSEIAPIS-----EEELERIVLA-----MHIHQONSHMETKTRKDP 832
Db 797 ---LTTKELKSDCAPLAGGEDGCAEDFDEAKLAQAREMIRYMALVSAMERTEHKARKG 852
QY 833 KDSIL----TRAQNMLRKRSYGNLKNVAESVEM 861
Db 853 TSALLSAKVIDAVMRKARPRYSPLRDTDBEEL 885

RESULT 28
VGBE31
Glycoprotein B - human herpesvirus 3
N:Alternate names: Glycoprotein II
C:Species: human herpesvirus 3, varicella-zoster virus
C:Date: 30-Sep-1988 #sequence_revision 30-Sep-1988 #text_change 16-Jul-1999
C:Accession: E27214
R:Davidson, A.J.; Scott, J.E.
J. Gen. Virol. 67, 1759-1816, 1986
A:Title: The complete DNA sequence of varicella-zoster virus.
A:Reference number: A27345; MUID:86306657; PMID:3018124
A:Accession: E27214
A:Molecule type: DNA
A:Residues: 1-868 <DAV>
A:Cross-references: EMBL:X04370; NID:G59989; PIDN:CAA27914.1; PID:G60020
C:Genetics:
```


Db 537 DMSR-MPGDPTMCYTRPVLIFRYSSSPESQFSANSTENHNLDILCOLGEHNEILOQRNLI 599
 Qy 604 ETCQENTYYFQAKTDMYIYKNYBHLKTVPLSSITLTDTFIALNFTLLENVDKFIELYT 663
 Db 596 EPCMINHRRYFLGNGVLLYEDYTFVRQVNASEIEEVSIFINLNATILEDLDVFPVEVYT 655
 Qy 664 RDEKRLSNVPDIETMFREYNYAOR-----VSLGRKDLLDLSTNRNQVDAFGSLMDD 716
 Db 656 REELRDGTGLNYDDVARYQNYNKRFRDITVIRGDRGDAI-----FRAIADFFGNTLGE 710
 Qy 717 LG-AVGQTVNNAVSGVATLFSISVTGFINFKNPPGGM-----LMIIVVIGVLFAIYFLT 770
 Db 711 VGRALGTVWTTAAAVI-----STVSGTASFISNPPAALGIGIAVVVSIILGLAFKPYMN 766
 Qy 771 KTKIYETAPIKMIYP 786
 Db 767 LKSN-----PQVQLFP 777

 RESULT 31
 VGBE2H
 glycoprotein B precursor - equine herpesvirus 1 (isolate HVS 25A)
 C:Species: equine herpesvirus 1
 C:Date: 30-Jun-1990 #sequence_revision 30-Jun-1990 #text_change 05-Jan-1996
 C:Accession: A31241; JQ0008
 R:Walley, J.M.; Robertson, G.R.; Scott, N.A.; Hudson, G.C.; Bell, C.W.; Woodw
 J. Gen. Virol. 70, 383-394, 1989
 A:Title: Identification and nucleotide sequence of a gene in equine herpesviru
 A:Reference number: A31241; MUID:89279217; PMID:2543744
 A:Accession: A31241
 A:Molecule type: DNA
 A:Residues: 1-980 <MHA>
 C:Superfamily: herpesvirus glycoprotein B
 C:Keywords: glycoprotein; transmembrane protein
 F:1-85/Domain: signal sequence #status predicted <SIG>
 F:70-78/Domain: transmembrane #status predicted <TN1>
 F:86-980/Product: glycoprotein B #status predicted <MAT>
 F:853-868/Domain: transmembrane #status predicted <TN2>
 F:165,275,380,423,497,514,515,560,727,749,952,971/Binding site: carbohydrate (

[illegible]

RESULT 32

B48349
glycoprotein B precursor - ateline herpesvirus 1 (strain Lennette)
C|Species: ateline herpesvirus 1
C|Date: 17-Feb-1994 #sequence_revision 17-Feb-1994 #text_change 09-Sep-1994
C|Accession: B48349
R|Eberle, R.; Black, D.
Arch. Virol. 129, 167-182, 1993
A|Title: Sequence analysis of herpes simplex virus gB gene homologs of two platyrrhine
A|Reference number: B48349; MUID:93228440; PMID:8385913
A|Accession: B48349
A|Molecule type: DNA
A|Residues: 1-933 <EBE>
A|Note: sequence extracted from NCBI backbone (NCBIN:129063, NCBIPI:129065)
C|Superfamily: herpesvirus glycoprotein B
C|Keywords: glycoprotein; transmembrane protein
F|1-35/Domain: signal sequence #status predicted <SIG>
F|36-933/Product: glycoprotein B #status predicted <MAT>
F|775-794/Domain: transmembrane #status predicted <TM1>
F|801-818/Domain: transmembrane #status predicted <TM2>
F|107_161_418_450_697_747/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 18.1%; Score 822; DB 1; Length 933;
Best Local Similarity 26.4%; Pred. NO. 6.6e-41;
Matches 224; Conservative 164; Mismatches 378; Indels. 84; Gaps 20

Qy	30	TGVTSPNTATWTSPELTGYTHGTHDSHGHERGNN-----ENRDEEONKNIGYSPSTFFY	85
Db	80	SGPESP-----GPDPRPGGGSGRRGSPGNGTRSAARRQLRESLRIRIQEYAAAF	133
Qy	86	RVCASGVGDVRFQTDHVCPD-ASDWHSGLLIYKQNIIPMFVRVKYRKVTTSTV	144
Db	134	YVCPPTGTATVQFSEPRPCPDVAAGKNFTEGIAVWFKNIAPKFTATKYKEITVSQT	193
Qy	145	Y-----NGIYSUSINQHTFFKSIIEPWE--TEKMDTYQCENSLRLTGGNLITYVD	194
Db	194	WQSGRYLOUTGLYND-----RAPVPFEIIDLINGKRCRSDVYTRSORRYTAVD	244

QY	195	RDDINMTYFLQVDPGVTDVKRYGQSQPELYLEPGFWGSYRRRTTVCNCLMDMFARSNPP	254
Db	245	GDWGREVALVPAKTSTSPNSRGWYTTDRVY-APNAHAGFYKTGTTVNCIVVEWEARSAPP	303
QY	255	PDFFTATGDTVEVSPFWSGEDD--HENKMHEKPMFVSVINNY-KVUDYQNGRTVPLGKT	311
Db	304	YDSFVLATGEFVYASPSFGSEDARRERNRYAPDRFQVDGVPYPRDLSGQRAATPV--V	361
QY	312	RIFLDREBYTSLWEKHLKNMYSVCLTLKAFYNGIQTEHSGSYHYFVANDITASFTTSKED	371
Db	362	RNLLTPTFTVGWDWKPRPNVCSVTKKQVVEEMVRAEYGSARFTTSAALSATFTS----	417
QY	372	MKEFNTTYH-----CLNBEIKAEIEKKYA-KVNSTHSKYGLKYPKTDDGGLVLMQ	421
Db	418	---NLQYPPELTEHSDCVAREAAEIEAIVARYYNASHVRVGVQVYLAAGGFLLAQ	473
QY	422	PLIQNRLLDAKNLNNETYSRRSRQAESTTDDPMEMTGNAG--GEVSSSENSITVAQVQ	479
Db	474	PLLSNSLAEMYRR---EALLGRSGDLAAALAPPVVAAPASGAGPRGTISTTQTVEFARLQ	530
QY	480	YAYDNLRIRINILEDLSKACREOHRALVKNELSKINPTSVMSIYNRPYSYAKRIGDV	539
Db	531	FTYDHIQKHVNEMLGRIAAACQONQELVLWNEARKLNPGAISATVGTVRGARMLGDV	590
QY	540	ISVNCIIVDQTSVSLKSLRLLSASDBKFSRPPVTTFKMNDDSTIYKQGLGVNNEILLT	599
Db	591	MAVSTCIPVSPDNVIMQNSMR-IPGDKTCYARPLVFSFRYTDDEGELVEQLGEDNEIRLE	649
QY	600	TTYLETQENTEYFFQAKTDMVIKYNIEHLKTVPLUSSITTLDTPTALNPTLELVDFPKVI	659
Db	650	QNNVEPCTVGHKRYFVFGGFPVYFEEYAYSHQVSADVPVSTFVDLNLTMLEDHEFLPL	709
QY	660	ELYTRDEKRLSNVDFIETMFREYNYAORVSGLRKDLDLSTNRRQVDAFGSLMDDLGA	719
Db	710	EYVTRREIKDGLDYAEVQRNQLHARLFSDDIRIMND-SANA-ALMAGLARFDFGMDG	767
QY	720	VGQTVNAVSGVATLFFSGIVTGFINFIKNPGFGLMIIVIGVLPAIYPLTKTKTIYETA	779
Db	768	AGKAIGRAVLGVTEGLISWGSVSSFLSPNPGFALAVGLVLVAGLVAAFFAMRYIMRLRAN	827
QY	780	PIKMIYPEI-----DKLKREGEKSEIAPISEEELE-----RIVLAMHI	817
Db	828	PMRALYPLTTSIGIKAEAPAAALGSGDKGAGGAGVEDFDEAKLEAARDMTRYMTLVA	887
QY	818	HQONSHMETK	827
Db	888	MERTAHKAKK	897
RESULT 33			
VGBEOH			
glycoprotein B precursor - equine herpesvirus 4 (strain 1942)			
C;Species: equine herpesvirus 4			
C;Date: 31-Mar-1990 #sequence_revision 31-Mar-1990 #text_change 16-Jul-1999			
C;Accession: A31880			
R;Riggio, M.P.; Cullinan, A.A.; Onions, D.E.			

RESULT 33

glycoprotein B precursor - equine herpesvirus 4 (strain 1942)
C:Species: equine herpesvirus 4
C:Date: 31-Mar-1990 #sequence_revision 31-Mar-1990 #text_change 16-Jul-1999
C:Accession: A31880
F:R:9910, M.P.: Cullinane, A.A.; Onions, D.E.
J. Virol. 63, 1123-1133, 1989
A:Title: Identification and nucleotide sequence of the glycoprotein gB gene
A:Reference number: A31880; MUID:89125704; PMID:2915378

Query Match	18.0%;	Score 819.5;	DB 1;	Length 919;
Best Local Similarity	26.5%;	Pred. No. 9.1e-41;		

#5

Db 796 DNTAVIMOGIATPFGKLGVEAVGTVLGAGAVVSTVSGIAFINNPGGLAIGLVI 855
Qy 761 GVLFAIYFLKTKIYETAPIKMIYPEI-----DKLERGCKSEIAPISE-----BELE 809
Db 856 AGLVAAPFAFYVNMQLRSNPMKALYPITTSRLKNKAKASGYQNDSDTSDFDAKUEEAR 915
Qy 810 RIVLAMHIHQNSHMETKTRDKP-DSILTRAQNMUKRSQ--YSNLKNAESVE 860
Db 916 EMIKYMSVSALEKQEKANKKNGKVGGLIASNVSKIALRRGPKYTLREDDPWE 970
RESULT 35
VGBBHH
glycoprotein B precursor - bovine herpesvirus 2 (strain BMV)
C;Species: bovine herpesvirus 2
C;Date: 31-Mar-1990 #sequence_revision 31-Mar-1990 #text_change 16-Jul-1999
C;Accession: C29242
R;Hammerschmidt, W.; Conraths, F.; Mankertz, J.; Pauli, G.; Ludwig, H.; Buhk, H.J.
Virolgy 165, 388-405, 1988
A;Title: Conservation of a gene cluster including glycoprotein B in bovine herpesvirus 2
A;Reference number: A94381; MUID:88306231; PMID:2841793
A;Accession: C29242
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 1-917 <HAM>
A;Cross-references: GB:M21628; NID:g330752; PIDN:AAA46053.1; PID:g330753
C;Superfamily: herpesvirus glycoprotein B
C;Keywords: glycoprotein; transmembrane protein
F;1-22/Domain: signal sequence #status predicted <SIG>
F;23-917/Product: glycoprotein B #status predicted <GPB>
F;578-594/Domain: transmembrane #status predicted <TM1>
F;770-786/Domain: transmembrane #status predicted <TM2>
F;795-811/Domain: transmembrane #status predicted <TM3>
F;48,110,164,278,421,453,505,564,692/Binding site: carbohydrate (Asn) #status
Query Match 17.7%; Score 805; DB 1; Length 917;
Best Local Similarity 27.5%; Pred. No. 6.6e-40;
Matches 229; Conservative 164; Mismatches 335; Indels 106; Gaps 27;
Qy 74 KNIYGSPTPYRVCSASGVGVDFRFTQDTHVCPD-ASDMVHSGILLIYKQNIIPMFRV 132
Db 125 RGLKSGDDPNFYVPPPTGATVVRLEPRPCPLKGLNFTGEGIAVTFKENLAPYKFA 184
Qy 133 RYRKVTTTGVNGIYSDITNQHTFYKSE-----PWETEXMDTIY---QCENSRLN 184
Db 185 TWYKAVTASVWSG-VS-----YNOFMNIFEDRAPIPFE-ELVDRIHGRGMCSTAK-- 235
Qy 185 TGGNLLTYVDRDDINMTVFLQPDGVTPDVVKRYGQPELYLEPGWF-----WG 232
Db 236 -----YV-RNNLETTAFNDAD--EHEMKLVPAESAPGLHGWHTTRLKNNPTGSAAW- 284
Qy 233 SYRRRTTVNCELMDMFARSNPPDFVTATGDTVEMSPFWSGEDDHENKMKHKEPWFVSVI 292
Db 285 IHRHGTIVDCTIVEAKSYNYNEFVLATGDFVYAPFF-GYRDGSHSEHN-----AYAA 339
Qy 293 NNKVVQYQNRGTVP---LGK-----TRIFLDREYTLSEKHLKNMYSYCPILTWKAF 342
Db 340 DRFKQVD-----GFFRDFGTRHRGSPVTYNLLTPTMTVGMWAPKRPVCTMTKREV 395
Qy 343 YNGIQTEHSGSYHFVANDITASFTT-----SKEDMKFENTTYHCLNEBIKAEIKKY- 394
Db 396 PEMLRABYSGSFRFTSNALSATFTTNLTQYSLSRVDLGD-----CVGKEAREADRIYL 449
Qy 395 AKVNSTHSGYDLKYFTKDGGLYLWQPLIQNRLLDANK- LNNETYSRRSRQASSTTD 453
Db 450 EKYNNTHLRVSGVQYLYATGGFLIAYQPLLSNNLADLYVKELMREQALKPEERKLNATD 509
Qy 454 PMEMTNGAGGESSENSITVAQVOYADNLRIRINNILEDLSKACROHRAALVWNE 513
Db 510 -----GKVTITSSVEFLRQTYNHIQKHVNMFGRAVMSWCELOQBELTWNE 559
Qy 514 LSKINPTSVMSIYNRPVSARIGDVISVSNCTIWDQTSVLSHKSRLLSASDEKCFSRP 573

Db 560 AKKINPSAIVTLHRRVSGACMLGDLVLAISTCVAVPAENVIMQNSMRIPS-KPGTCYSRP 618
Qy 574 PVTFRKWNDSITVKGOLGVNNEILLTTTILETCQENTTEYYFOAKTDWYIYKNVHLKTVP 633
Db 619 LJSFKHVDGEELMEGOLGNNELRLDRDAVEPCVSGHKRYFLFGAGVYFEEYTSHQLS 678
Qy 634 LSSITLDTFIALNFTLLENVDPKVELYTRDRSKLSNVFDIETMPREYNYAQRVSGLR 693
Db 679 RSDITAVSTFIDLNTIMLEDHEFPVLEVTRQBIKDSGLDYAEVORRNOLHARADI- 737
Qy 694 KOLLDLSTNRNQFVDAFGSLMDDLGAVGQTVNNAVSGVATLFSIVTGFNFINKNPFGM 753
Db 738 DTVIKADPNAAI FAGLHG-FFEGLDGVDGRAVGWLGVGVVATVSGVSFLSNPPGAL 796
Qy 754 LMIIVIGVLFALYFLTKTKIYETAPIKMIYPEIDK-LKERE-----GKSEIAPISEEL 808
Db 797 AIGLLVLGGLVAFAAFYVRLQRPNMKALYPLTTKDLKHPSEGGGGEAMDFDEQKL 856
Qy 809 ERIVLAMHIHQNSHMETKTRDKPDSILTRAQNMUKRSQYSLNKNVAESVEML 862
Db 857 DEARSMIKYVALVSAME-----RTYKHAGRRGGTSAILNARLTDNV 897
RESULT 36
B48474
glycoprotein B - feline herpesvirus 1 (fragment)
C;Species: feline herpesvirus 1
C;Date: 07-Apr-1994 #sequence_revision 18-Nov-1994 #text_change 26-Aug-1999
C;Accession: B48474
R;Spatz, S.J.; Maes, R.K.
Virology 197, 125-136, 1993
A;Title: Immunological characterization of the feline herpesvirus-1 glycoprotein B and
A;Reference number: A48474; MUID:94025559; PMID:8212548
A;Contents: C-27
A;Accession: B48474
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-943 <SPA>
A;Cross-references: GB:S63371; NID:g435792; PIDN:AAB28559.1; PID:g435794
A;Note: sequence extracted from NCBI backbone (NCBIN:138805, NCBIPI:138807)
C;Superfamily: herpesvirus glycoprotein B
C;Keywords: glycoprotein
Query Match 17.4%; Score 789.5; DB 2; Length 943;
Best Local Similarity 27.3%; Pred. No. 5.7e-39;
Matches 231; Conservative 151; Mismatches 364; Indels 101; Gaps 27;
Qy 80 PSTFPYRVCSASGVGVDFRFTQDTHVCPDAS-DMVHSEGLIYKQNIIPMFRVRYKRV 138
Db 121 PSTF--YMCPPSPSGSTVVRLEPPRACPDYKLGKNFTGEGIAVIFKENIAPYKFKANIYKN 178
Qy 139 VTTSTVYNGIYSDITNQHTFYKSEIPEWE-TEKMDTIYOCFNSRLNLTGNNLLTYVDRDD 197
Db 179 IIMTTVMWSSGYAVTTNRYTDRVPVKVQBITDLIDRRGMCLSKADYVVRNYYQTFADRDE 238
Qy 198 INMTVLOQVDGVTDPVKYGSQPELYLEPGWFGS-----YRRRTTVNCELMDM 247
Db 239 DRPELPLKPPSTLSRVR-----GHTNNTYTKIVLLDFHSGTSCVNCIVEEV 286
Qy 248 FARNSNPPDFVTATGDTVEMSPFWSGED-----DHENKMKHKEPWFVSINNYKVVDYQNR 303
Db 287 DARSVYVYDPSFALSTGDIHMSPPFFGLRGDAHVHETSYSDDR--FQIEGYYPIDLDTD 343
Qy 304 GT-VPLGKTRI FLDRREYTLSEKHLKNMYSYCPILTWKAFYNGIQTEHSGSYHFVANDIT 362
Db 344 YTGAPV--SRNLETPHVTAVMWTPKSGRVCTLAKWREI-DEMLPMNITGSGYRFTAKTIS 400
Qy 363 ASFTTSKEDMKFENTTY--HCLNEEIKAEIKKY-AKVNSTHSGYDLKYFTKDGGLYL 419
Db 401 AFIISMSTQF-EINRIRLGDCAKAAEAIDRIYKSKYKTHIQTGTLTYLARGGFLIA 459
Qy 420 WOPLIQNRL-----LDANKLN--NETYSRRSR-----ROAESTTDP 454

Db 460 FRPMISNELAKUYINELARSNRTYVLDLSALLNPSGETVQTRRSVPSNQHRRSRSTIEG 519
QY 455 MMEMTGAGGYSSENSITVAOVAYDNLIRININLELDSKAWCREOHEAALVWNEEL 514
Db 520 GIETVNNAS--LLKTTSSVEPAMQFAYDIOAHVNEMLSRITATWCTLQNHVIMTET 577
QY 515 SKINTPSVMSMIYNRPVSAKRIGDVISNSCIVDPQTSVSLHKSURLLSASDEKCFSRPP 574
Db 578 LKLNPGGVVSMALERRVSARLLGDAAVAVTQCUNISSGHVYIQNSMR-VTGSSTTCYSRPL 636
QY 575 VTFKFWNDSTIKYKOLGVNNEILLTTTLYLETQENTYFYQAKTDMYIKYKNEHLKTVPL 634
Db 637 VSFALNDSEXTEGOLGENNELVERKLEPCTVANKRYFKFGADYVVFEDYAYYRKVPL 696
QY 635 SSITLTDIFIALNFTLLENVDFKVIELYTRDEKRLSNVFDIETMREYNYAORVSGLRK 694
Db 697 SEIELISAVY-IKSTILLEDREF-LHSSYTRALEBDTGPDYSEIORRQLHALKFYDI-- 752
QY 695 DLLDLSTNRNQFVDAFGSLMDLGAQGVQTVVNAVSGVATLFSSIVTGFINFIPKPFPGMW 754
Db 753 DSIVRVDDNVLIMRGWANFFOGLGDVGAGFGKVLGAASAVISTVGVSSFLNPFGLA 812
QY 755 MIIVIGVLPALYFLTKTKIYETAPIKMIYEPIDKLKEREKSE-----IAPI 803
Db 813 VGLLLAGIVAAFLAYRYSRLRANPMKALYPTVTRNLKQTAQSPASTAGSDSPGVDDF 872
QY 804 SEELER-----IVLAMIHOONSHMETKTRKDPKDSILT-RAQNMLRKRS--YS 851
Db 873 DEELMQAREMIKYMSLVSM---EQECHKAMKNGP--ALTSHLTNMLARRRGPYQ 927
QY 852 NLKNAES 858
Db 928 RLNNLDS 934

RESULT 37
JH0109
glycoprotein 14 precursor - equine herpesvirus 1
C:Species: equine herpesvirus 1
C>Date: 31-Dec-1991 #sequence_revision 31-Dec-1991 #text_change 26-Aug-1999
C:Accession: JH0109
R:Guo, P.
Gene 87, 249-255, 1990
A:Title: Characterization of the gene and an antigenic determinant of equine herpesvirus
A:Reference number: JH0109; MUID:90236317; PMID:1692002
A:Accession: JH0109
A:Molecule type: DNA
A:Residues: 1-979 <GUO>
A:Cross-references: GB:M34861; NID:g330900; PIDN:AAA46086.1; PID:g330901
A:Note: glycoprotein 14 is homologous to corresponding glycoprotein of pseudorabies, bo
C:Genetics:
A:Gene: gp14
C:Superfamily: herpesvirus glycoprotein B
C:Keywords: glycoprotein; transmembrane protein
F:1-86/Domain: signal sequence #status predicted <SIG>
F:87-979/Product: glycoprotein 14 #status predicted <MAT>
F:831-871/Region: hydrophobic
F:165,275,380,423,497,514,560,727,749,951,970/Binding site: carbohydrate (Asn) (covalent

Query Match 17.1%; Score 777.5; DB 2; Length 979;
Best Local Similarity 26.3%; Pred. No. 3.1e-38;
Matches 247; Conservative 159; Mismatches 377; Indels 155; Gaps 34;

QY 32 VTSPNWTATWSTESPLTHGYTHDSSHGREGNENRDSEONKNYIGSPSTFPYRVC-SA 90
Db 83 VRAVPTTSPPTSTPS--MSTH--SHG-----TVDPDLLPTETDPLRLAVRE 127
QY 91 SCV-----GD-----VRFQTDHVCDDAS--DMVHSEGLLIYKQNIIPFVRXK 134
Db 128 SGILEADGDFTYCPPTGTVRIEPPPTCPKFDGRNFTGIAVIFKNIAPYKFRANV 187
QY 135 YRKVVTSTVNGIYSDSITNOHTFYKSIPEWTEK---MDTIYQCFNS---LRLNTGG 187

Db 188 YYKDIVTVRWKGYSHTSLSDR---YNDRVPSVVEEIFGLIDSKGCKSSKAEYLRL---D 240
QY 188 NLLTYV---DRDDINMTVFLQVDPVDPVKRYGSOQELYLEPGWF-WGSYRRRTUNCE 243
Db 241 NIMHAYHDDDEVELD--LVPSKPAFGARAWQTTNDTTSYVGMPWRHY-TTSVNCI 297
QY 244 LMDFAIRSNPPDFEFTATGTVMSFPW----SCEDDHENKMKHEKPMFVSVINNYKVD 299
Db 298 VEEVEARSVPYDVSFALSTGDIVVASPFYGLRAAAIEHNSYAQER---PRQVEGYRRPD 354
QY 300 YONRGTVPLGKTRIFLDRREEVTLNWEKHLKMSVCPILTKAFYNGIOTEGSHSGSVHFVAN 359
Db 355 LDSLQAEETPKNFITTPHVTVSNWTEKKEVEACTLTWKKEVDELVDLDEFRGSYRFTIR 414
QY 360 DITASFTTSKEDMK-EFNTTYHCLNEELKAEIKKYAK-VNSTHSKYGLKAYFKTDGGLY 417
Db 415 SISSYFISNTTQFKDESAPLTECVSKEAKEAIDSIIYKQYESTHVFSGDVYIYLRGGFL 474
QY 418 LVWQPLI-----QNRLLDAKNKL-----NNETYSRRSRROAESTTDPMMEMT 459
Db 475 IAPRMLSNELARLYLNLVRSNRTYDLKNLLNPNANNNTTTRRRSLLSVPEPQPTQD 534
QY 460 G-----NGAGGEYSSE-----NSITVAQOVAYVNLRIIRINI 492
Db 535 GVHREQILHLRHKRAVEATAGTSSNVTAKQLELIKTTSSIEFAMLQPAYDHIQSHVNE 594
QY 493 LEDLSKACREOHRALVWNELSKINPTSVMSMIYNRPVSAKRIGDVLVSNCIVDQTS 552
Db 595 LSRIATACWPLQNKERPLWNEVMKITPSAIVSATLDERVAARVLGDVIAITHCAKI-EG 653
QY 553 VSLHKSURLLSASDEKCFSRPPVTF---KFWNDSTIKYKOLGVNNEILLTTTLYLETQEN 609
Db 654 VYLQNSMR--SMDSNTCYSRPPVTFITKQANNRSGLEGOLCEENEIFTERKLEPCALN 711
QY 610 TEYFQAKTDMYIKYKNEHLKTVPLSSITLDTFIALNFTLLENVDFKVIELYTRDEKRL 669
Db 712 QKRYFKGKEYVYENYTVFRKVPPTIEVISTVVELNLTLEDREFLPLESVYTRAELED 771
QY 670 SNVFDIETMFREYNYAORVSGRLKDLLDLSTNRNQFVD-----AFGSLMDDLGA 719
Db 772 TGLLDYSEIQRRNLHALLR-----FYDIDSVVN--VDNTAVIIRGSPAFSAWVKWR 822
QY 720 VQOTVNAVSGVATLFSSIVTGFINFIPKPFPGMWLIIVIGVLPALYFLTKKIYETA 779
Db 823 PWERSFGARGAV---STVSGIACFLNPFGLAIGLLVIAGLVAFAFFAYRYVMQIRSN 878
QY 780 PIKMIYPEIDKLKEREKSEIAPISE-----EELERIVLAMHIHQNSHMETKT 828
Db 879 PMKALYPITTKALKNKAQTSYQNEEDDGSDFDEAKLEAEAREMIKYMSVMSVALEKQEK 938
QY 829 RKDPKDS-----ILTRAQNMLRKRS--YSLNKAESVE 860
Db 939 IK--KNSGVGLIASNKLALRRRGPKYTRLQONDITME 974

RESULT 38
VGBEPS
glycoprotein gII precursor - suid herpesvirus 1
C:Species: suid herpesvirus 1
C>Date: 31-Mar-1988 #sequence_revision 31-Mar-1988 #text_change 16-Jul-1999
C:Accession: A29159
R:Robbins, A.K.; Dorney, D.J.; Wathen, M.W.; Whealy, M.E.; Gold, C.; Watson, R.J.; Holla
J. Virol. 61, 2691-2701, 1987
A:Title: The pseudorabies virus gII gene is closely related to the gB glycoprotein gene
A:Reference number: A29159; MUID:87284141; PMID:3039163
C:Accession: A29159
A:Molecule type: DNA
A:Residues: 1-913 <ROB>
A:Cross-references: GB:M17321; NID:g334053; PIDN:AAA47465.1; PID:g334054
A:Note: the authors translated the codon GAC for residue 860 as Asn
C:Superfamily: herpesvirus glycoprotein B
C:Keywords: glycoprotein; transmembrane protein
F:1-40/Domain: signal sequence #status predicted <SIG>

F;751-819/Domain: transmembrane #status predicted <TM>
F;820-913/Domain: intracellular #status predicted <INT>
F;151,261,441,516,633,697/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 17.18; Score 775.5; DB 1; Length 913;
Best Local Similarity 27.08; Pred. No. 3.7e-38;
Matches 235; Conservative 165; Mismatches 388; Indels 81; Gaps 26;

Qy 26 SIATGVTSPNTATWTEPLCHYTHDSHGCGNNENRSEEQNKNIYSPSTFPY 85
Db SLEIEAFSP-----GPSEAP-DGEYGLDARTAVRAAATERDR-----F 123
Qy 86 RVCSASGVDFRQTDHVCDDAD-MVHSEGLLIYKONIIPMPFRVRYKRVTTSTV 144
Db YVCPPPSGSTVRLPEPQACPEYSQGRNFTTEGTAIVLAKFKAHYIKYKNIIVTV 193
Qy 145 YNGIYSDSIHQHTFYKSIPEWE-TEKMDTIYQCFNSRLNTGNNLLTYVDRDDINMTVF 203
Db WSGSTYAAITNRFTDRVPVPVQEIITVDIDRRGCVSKAEYVRNNHKVTAFTDRDENPVEVD 243
Qy 204 LQPDVGVTPDKVYSGOPELYLEPGFWGYSRRRTTVNCELMDMFARSNPPFFVATG 263
Db LRPSRLNALGTRGHHNTDITYKIG-AAGFYHTGTSVNCIIVEVEARSVPYDPSFALSTG 302
Qy 264 DTVMSPFWS-GEDDHENKHEKHPFVSVVINNYKVVDYQNRGTVPLGKTRIFLDREYTL 322
Db DIVMSPFYGLREGAHEHGYAPGRQOEYHYPIDLSRLRASESVTRNFTPHFTV 362
Qy 323 SWEKHLNMSYCPPLTLWKAFYNGIQTE-HSGSYHFVANDITASFTT--SKEDMKEFNITY 379
Db AWDWAPTRRVCSLAKWREAEEMTRDETROGSRFTSRALGASVSDVTQLDQRLVHLG- 421
Qy 380 HCLNEEKABIEKKY-AKVNSTHSGYDGL-KYFKTDGGLYLWQPLIQNLRLDA-KNKLN 436
Db DCVLRASEAIDATYRRYRNSHTVLADRPVYLARGFVVFAPRPLSNELAQIYARELE 481
Qy 437 N-----ETYSRRSRQ--AESTDPMEMTNGAGGEVSSSENSITVAQOVAYD 483
Db RLGLAGVVGPAPAAARRARRSPGAGTPEP---PAVNGTGHURITTGSAEFARLQTYD 538
Qy 484 NLRIIRINILEDLSKACREQHRALVYNELSKINPTSVMSMIYNNRPVSAKRIGDVISVS 543
Db HIQAHVNDMLGRIAAACELQNKDRTLWSEMSRLNPSAVATAALGQSVASARMGLGDVMAIS 598
Qy 544 NCIVVDOTSVLHKSLLSASDEKCFSPRPVTFKFMNDSTIYKQGLGVNNEILLTYYL 603
Db RCVEV-RGGVYVQNSMR-VPGERTCYSRPLVTFEH-NGTGVIEGQLGDDNELLISRDLI 655
Qy 604 ETQCENTEYFQAKTDMVIYKVEHLKTVPLSSITLDTFIALNFTLLENVDFKVIELYT 663
Db EPTGNHRRYFKLGSVYVYEDYVYVRNVEPVE--TISTRTVNLTLLEDFLEPLVEYIT 713
Qy 664 RDEKRLSNVFDIETMFREYNYYA-----QRVSGLRKDLDDLSTNRNQFVDAFGSLMDDL 717
Db REELADTGLLDYSEIQRRNQLHALKFYDIDRVKVDHNVLLRGIANFF-----QGL 765
Qy 718 GAVGQTVNAVSGVATLFLSSIVTGFINPKNPPGCMMLIIVGLVPAIYFLFKTKIYE 777
Db GDVGAAYGVKVLGATGAVISAVGMVGSFLNPGFALAIGLLVLAGLVAAPFLAYRHISRLR 825
Qy 778 TAPTKMIIPEIDKLKREGKSEIAPISEBELER-----IVLAMHIIHQONSHMETKT 828
Db RNPWKALYPVTTKLEGGVDE-GDVDEAKLDQARDMIRYMSIVSAL---EQGEHKARKK 881
Qy 829 RKDPKDSILTRAQNMRLKRSYSLKNNAE 857
Db NSGPA-LLASRVGAWATRRHYQRLSESD 909

RESULT 39
VGEBEG
glycoprotein g1 precursor - bovine herpesvirus 1
N;Alternate names: glycoprotein 11a; glycoprotein 16; glycoprotein g130; glycoprotein gV

C;Species: bovine herpesvirus 1
C;Date: 31-Mar-1990 #sequence_revision 31-Mar-1990 #text_change 16-Jul-1999
C;Accession: A31166
R;Misra, V.; Nelson, R.; Smith, M.
Virology 166, 542-549, 1988
A;Title: Sequence of a bovine herpesvirus type-1 glycoprotein gene that is homologous to
A;Reference number: A31166; MUID:89020821; PMID:2845660
A;Accession: A31166
A;Molecule type: DNA
A;Residues: 1-928 <MS>
A;Cross-references: GB:M23257; NID:G340858; PIDN:AAA46013.1; PID:G511852
A;Note: the authors translated the codon CTG for residue 534 as Ser
C;Superfamily: herpesvirus glycoprotein B
C;Keywords: glycoprotein; transmembrane protein
F;67/Domain: signal sequence #status predicted <SIG>
F;68-928/Product: glycoprotein g1 #status predicted <GGI>
F;756-824/Domain: transmembrane #status predicted <TM>
F;105,153,442,484,579,637,703/Binding site: carbohydrate (Asn) (covalent) #status predict

Query Match 15.9%; Score 722; DB 1; Length 928;
Best Local Similarity 26.8%; Pred. No. 5.7e-35;
Matches 240; Conservative 156; Mismatches 427; Indels 94; Gaps 26;

Qy 2 AGSLKLRGSLALWLYLQVALYSLSAETGVT-----SPPNTATWSTESPLTGHY 51
Db 41 AGGARAALAAALLMATWALLLAAPAGRAPATTPAPPPEEAASAPPASPSPGPDGDDA 100
Qy 52 GTHDSHGGRGNNENRDSSEQNKNIYGPSPTFPYRVCSASGVGDVFRFQTDHVCDD-ASD 110
Db 101 ASPDNDTDVRAALRLAQAAGENSRRF-----VCPPPSGATVVLAPARPCPEYGLG 151
Qy 111 MVHSEGLLIYKONIIPMPFRVRYKRVTTSTVYNGIYSDSIHQHTFYKSIPEWE-TE 169
Db 152 RNYTEGIVYKENTIAPTYTKAIYYKNIYVITTWAGSTYAAITNQYTRVPVGMGEITD 211
Qy 170 KMDTIYQCFNSRLNTGNNLLTVVDRDDINMTVFVLPDVGTVPDVKRYGSOPELYLEPGW 229
Db 212 LVDKKWRCLSKAEYLRSGRKVAVAFDDEDDPWEAPLKPALSAFGVGRWHTDDVYTAG- 270
Qy 230 FWGSYRRRTTVNCELMDMFARSNPPDFVTFATGDTVMSPFWS-GEDDHENKHEKHPWF 288
Db 271 SAGLYRTGTSVNCIIVEARSVPYDPSFALSTGDIYYSPFYGLREGAHEHTSYSPER 330
Qy 289 VSVINNYKVVDYQNRGTVPLGKTRIFLDREEVTLSEKHLKNMSYCPPLTWKAFYNGIQT 348
Db 331 FOQIEGYKRDWATGRRLKEPVSRRFLRTOHVTVAWDVFKRKNVCSLAKWREADEMLRD 390
Qy 349 EHSGSYHFVANDITASFTTSKEDMKEFNITY-HCLNEEKABIEKKY-AKVNSTHSGYKD 406
Db 391 ESGNFRFTARLSATFVSDSHTFALQNVPLSDCVIEEAEAAVERVYRERYNGTHVLSGS 450
Qy 407 LKYFKTDGGLYLWQPLIQNLRLDAKNKL-----NNETYS-----RRSRQAE 449
Db 451 LETYLARGGVVAVAFRPMLSNEL-AKLYLQELARSNGTLEGLFAAAAPKPGPRARRPRR 508
Qy 450 SITDPMEMTNG-----AGGEYSSSENSITVAQOVAYDNLRIINILEDLSKAWC 501
Db 509 LRPAPE-----GRQARRRRHAGGRVTVSLAEFAALQTHDHTRTSEHHV-HRLASPPWC 562
Qy 502 REQHRAALVYNELSKINPTSVMSMIYNNRPVSAKRIG-----DVI--SVSNCIYVVDQTSVSL 555
Db 563 LIQNKERALLWAAAKLNPSAAASALDRPPRACWGTGHRDVLPRAGRALFIENSARA 622
Qy 556 HKSRLRLSASDEKCFSPRPVTFKFMNDSTIYKQGLGVNNEILLTYYTLETQENTYFYFQ 615
Db 623 RR--RL-----QPPRFLSFGNESEPEVQOLCEDNELLPGRELVEPCTANHKKRYFR 671
Qy 616 AKTDMVIYKVEHLKTVPLSSITLDTFIALNFTLLENVDFKVIELYTDEKELSNVFDI 675
Db 672 FGADYVYVYENYAVRRVPLAELEVISFTFDLNLTVLEDFLEPLVEYITRAELADTGLLDY 731
Qy 676 ETMFREYNYYAQRVSGLRKDLDDLSTNRNQFVDAFGSLMDDLGAVGQTVNAVSGVATLF 735

Db 732 SEIQRRNQJHEURFYDI--DRVVKTDGNMAIMRGLANFFQGLGAVGQAVGTVVLAGAGAA 789

QY 736 SSIVTGFINPKNPFQGMMLIIWIGVFAIYPLTKTKTIYETAPIKMIYPEIDK-LKER 794

Db 790 LSTVSGIASFIANPFQALATGLLVLAGVAFLAYRISRLRSNPKWALYPITTRALKDD 849

QY 795 EKKSEIAPISBEE-----LER---IVLAMIHQONSHMETKTRDKPKDS--ILTR-AQN 842

Db 850 PGRNR--PGESEEFDAKLEQAREMIKIMYSLVSAVERQEHAKKSKNAARLLATRLTQL 907

QY 843 MLRKES--CYSNLKNAE 857

Db 908 ALRRRAPPEYQQLPMAD 924

RESULT 40

S68553

surface layer protein tetraabrachion precursor - Staphylothermus marinus

C:Species: Staphylothermus marinus

C>Date: 27-Nov-1997 #sequence_revision 12-Dec-1997 #text_change 17-Mar-1999

C:Accession: S68553; S78081

R:Peters, J.; Baumeister, W.; Lupas, A.

J. Mol. Biol. 257, 1031-1041, 1996

A:Title: Hyperthermostable surface layer protein tetraabrachion from the archaeobacterium cure.

A:Reference number: S68553; MUID:96192087; PMID:8632466

A:Accession: S68553

A:Molecule type: DNA

A:Residues: 1-1524 <PET1>

A:Cross-references: EMBL:U57967

A:Experimental source: strain F1

A:Accession: S78081

A:Molecule type: protein

A:Residues: 40-45;91-95;151-158;327-331;445-449;490-497;499-505;596-614;617-624;638-648;5-1451 <PET2>

A:Experimental source: strain F1

C:Complex: heterotetramer; two alpha chains and two beta chains

C:Keywords: blocked amino end; glycoprotein; membrane bound

F:1-26/Domain: signal sequence #status predicted <SIG>

F:27-700/Product: surface layer protein tetraabrachion alpha chain #status predicted <MAT>

F:701-1524/Product: surface layer protein tetraabrachion beta chain #status experimental

F:1019-1029/Region: hinge

F:1495-1519/Domain: membrane anchor #status predicted <MBN>

F:444,605,641,685,708,1279,1402/Binding site: carbohydrate (Asn) (covalent) #status experimental

Query Match 3.6%; Score 162; DB 2; Length 1524;

Best Local Similarity 18.1%; Pred. No. 0.22;

Matches 148; Conservative 132; Mismatches 301; Indels 236; Gaps 35;

QY 24 SLISIAETGVTSPPTATWSTESPLTCHYGTGTHDSHGGRNNENRDSERQN-----KNYIGS 79

Db 729 TLSI--TVVGAAPN-----TEYNFTFGYQVHDLNIGITRISPOWNGTWNISLVTDIYGT 781

QY 80 PSTPYRVCSAGVGDFRFDHVCPPDASDMVHSEGIILYKQNIIPFMFRVKRKWV 139

Db 782 GST-----SVPLITLYPTSIVINATWDVI---TWLRLSGSGTDLDFS----- 822

QY 140 TTSTVYNGIYSDISITNQHTFYKSIPEWTERKMDTIYQCFSRLNTGGL-----LTYV 193

Db 823 --DVSYNG-FIDNLITPTIYVFG-----PSDITPGSGFN-IYVNTTYNVSVRVAVDYL 871

QY 194 DRDDINMTV--FLQPDVGVTDPVKRYGSOPELYLEP-----GWFGWSYRRRTTVN 241

Db 872 PRNVVISVPEVLPGDITVQIFPHHNEVWGFIETALFDENQLLGY-----LT 922

QY 242 CELMDMFARSNPPDFPFVATGDTVE--MSPFWSGE---DDHENKMKHPWFV----- 289

Db 923 VRLVDPLSNT-----VVERVAGYVAGNLIVEDVDGDGDNEVFWVNLTAFLV 969

QY 290 -SVINNYKV-----VDYQNRGTVPGLKTRIFLDREYTLSEKHLKNNMSYCPFLTWKAFY 343

Db 970 LGVDKTYRVDVFLFVLAVLNPSSNITGVTA--DNECYV---QLDLNGTIY-----WNLGL 1019

QY 344 NGIQTEHSGSYHFVANDITAGFTTTSKEDWKBFNTTYHCLNEEIKAEIEKKYAKVNSTHSHK 403

Db 1020 SGIMLGGDQIVTVLVGLEKLDITKDGIAEINATVNDINTYLVKNVTDLLKTINNS---- 1076

QY 404 YGDLKYFTKDGGLYLWQPLIQNRLLDAKNKLN--NETYS-----RSTRQA 448

Db 1077 ---VVMKNKNDTATLIIGQAEIKAKLDLNLTSQVNDQVTMILACCNNAKSVLNRMEGTL 1133

QY 449 ESTTDPNMTEMGAGGEYSSENSITVAQVQYVDNLRIRINNILEDLSEKAWCRQHRAA 508

Db 1134 NSTYTGVLNVKSD--LSTLIDTVNNVVIIPKFNELVDNVTVNEINASRD----- 1178

QY 509 LVNNELSKINP--TSVMSMIYNRPVSARKRIGDVISVNCIVVDQTSVSLHSLRLLSASD 566

Db 1179 LIIQKISSVNDSLTIIISAGFN-----DVEAMISNLTLLNLRID 1218

QY 567 EKCFSRPPVTPKFWNDSTIYKQGVNNEILLTYYLETQCENTEYFYQAKTDMYIK-- 624

Db 1219 E-----LEGTL-----LFYMTANEQRLEGIINETADDIVYRLT 1251

QY 625 ----NYEHLAKTVPLSSITLIDTFI-----ALNFTLLENVDKVIETLREDEK 667

Db 1252 VIIDRYESLKNLITLRADRLMIINDNVSTILASIGNVNLTIVFNKLNDEIELGDNVAT 1311

QY 668 RLSNVFDIETMFREYNYVAQRVSLRKLDDLLSTNRNQFVDAGSLMDDLGAVGTGVNA 727

Db 1312 INAGIFQISNLGNANQ-----LLDLTSSKVEILLNAISSNASSAISSEIHNA 1359

QY 728 VSGVATLFPSSI-----VTG-----FINFIKNPFGGM 753

Db 1360 VNQLSTVLVQVNDTLTKITGEADNILNFLSSLEGSM 1396

RESULT 41

T28238

ORF MSV077 hypothetical protein - Melanoplus sanguinipes entomopoxvirus

C:Species: Melanoplus sanguinipes entomopoxvirus

C>Date: 21-Jan-2000 #sequence_revision 21-Jan-2000 #text_change 21-Jul-2000

C:Accession: T28238

R:Afonso, C.L.; Tulman, E.R.; Lu, Z.; Oma, E.; Kutish, G.F.; Rock, D.L.

J. Virol. 73, 533-552, 1999

A:Title: The genome of Melanoplus sanguinipes entomopoxvirus.

A:Reference number: 220484; MUID:99102612; PMID:9847359

A:Accession: T28238

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-598 <AFO>

A:Cross-references: EMBL:AF063866; NID:g4049647; PIDN:AAC97632.1; PID:g4049672

C:Genetics:

A>Note: MSV077

Query Match 3.5%; Score 158.5; DB 2; Length 598;

Best Local Similarity 21.1%; Pred. No. 0.087;

Matches 106; Conservative 88; Mismatches 192; Indels 117; Gaps 28;

QY 318 BEYTLSEKHLKNMSY---CPLTLWKAFFNGIQTEHSGSYHFVANDITAGFTTTSKEDMK 373

Db 169 DEYSNIFKRCSNLKYEIGKILV-----NDLQVK---EYIWDN-----SHKFDIK 212

QY 374 ---EFNTTY-HCLNEEIKAEIEKKYAKVNSTHSHKYGDLKY-----FKTDGGL 416

Db 213 FSEQYFITPYFVSVNTGVLTKFENGFAELYSISSNQIIKYDHVITCKSDFINIENG- 271

QY 417 YLVWQPLIQNR-LDAKNKLN-NETYSRRRRRQAEETDPMEMTNGAGGEYSSENSIT 474

Db 272 YFISMPNEGQTVLSGRELNVWSQTFCKTCRNTEKHFVLNVTLANSKNEY-----IN 326

QY 475 VAQVQYAYD-NLRIRINNILEDLSEKAWCRQHRAAALVWNELSKINPTSVMSMIYNR-PVS 532

Db 327 IARFQEVLDKLLKIITKNL-----DTLCKNYVIYRKITYTEICKNNPOCYMKYFLNHYNVK 381

QY 533 AKRI-GDIVSVSNCIVVDQTSVSLHSLRLLSASDEKCFSRPPVTFK-----FNWDS 583

382 ARYGGNIGIKYICIEITDFEII-----IKOYNKYTKDCEIFVPMKININNTKFCGYMNP 436
Qy 584 TIYKQGLGVNNEILLTTTLETQENTYFQAKTDMYIKYVYEHKLTVPSSII--TTLD 641
Db 437 T-----NEVTFESPKNDYCSIT-YVDINDTLYISNNQINVSTOKIHTYGHETLL 486
Qy 642 TFIALNFTLNVDFKVIETRDEKRLSNVDFIETMFREYNYAQRVSGRLKDLDLST 701
Db 487 TNKISPIILKVIN--ISDLYF-EQSTFQDVYDI-----AYHFEDELINPKQTYNFTS- 537
Qy 702 NRNOFVDAFGSLMDDLGAAGVGTNNVAVSGVATLFSSTIVTGFNFKPFGMLMIIVIG 761
Db 538 -----SMDPLNFHG-----IEHIIIFIGFVIF-----GIFLVRYS 569
Qy 762 VLFA-IYFLTKTKK-IYETAPIK 782
Db 570 QIFNSCASIFRKTIVYENIPMK 592

RESULT 42
T28677
rhoptry protein - Plasmodium yoelii
C:Species: Plasmodium yoelii
C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 01-Dec-2000
C:Accession: T28677; C45521
R:Keen, J.; Sinha, K.; Brown, K.; Holder, A.
Mol. Biochem. Parasitol. 65, 171-177, 1994
A:Title: A gene coding for a high molecular mass rhothry protein of Plasmodium yoelii.
A:Reference number: 220508; PMID:95021522; PMID:7935623
A:Accession: T28677
A:Status: preliminary
A:Molecule type: DNA
A:Cross-references: translated from GB/EMBL/DBJ
A:Residues: 1-2269 <KEE>
A:Cross-references: EMBL:L27838; NID:9457145; PID:9457146; PIDN:AAA21304.1
R:Keen, J.; Holder, A.; Playfair, J.; Lockyer, M.; Lewis, A.
Mol. Biochem. Parasitol. 42, 241-246, 1990
A:Title: Identification of the gene for a Plasmodium yoelii rhothry protein. Multiple co
A:Reference number: A45521; MUID:91101660; PMID:2270106
A:Accession: C45521
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 2131-2269 <KE2>
A:Cross-references: GB:M34283

Query Match 3.4%; Score 154; DB 2; Length 2269;
Best Local Similarity 18.0%; Pred. No. 1.2;
Matches 151; Conservative 154; Mismatches 306; Indels 228; Gaps 41;
Qy 93 VGDVFRFQTDHVCDDSDMVHSEGLIYKQNIIPFMRVYKRVKVTSTVY--NGIYS 150
Db 91 IRELFFKYLD-----EEAERKYLEGLKLELNKKIKDIIAKIEYKNTVELKKEKNWAYI 146
Qy 151 DSITNQ-----HTFYKSIETPWEKMDTIYOCFNSRLNTGGLNLTIVVDRDDI 198
Db 147 DELANQSPYKVTGYENKNTIYTIKSY---PQIYE-----GDIIDFY--NEL 190
Qy 199 NMTVFLQPVQDVTDPVX-----YGSQPELYLEPCWFGWGSYRRRTTVNCELMDMF 248
Db 191 SSVIKEDPIDIEDKTKLENLRSKIDNVYDKIQKMEIE-----TVKSHLNIE 238
Qy 249 ARSNPPDFFTATGDTGVENSFPWSGE-----DDHENKMKHKEKPFVSVNNYKVV 299
Db 239 TNNKLP-----NTILEIKYIYDEISKELNKMLEDFFKNKE-----LSN-KISD 282
Qy 300 YQNGRTVPLGKTRIFLDEEYTLSEWHLKMSYCPULTLWKAFFN-----GIQTEHSGS 353
Db 283 YDKK-----REQLSYKSKMLEIRNH-----YNSQTVNDNTKEEBAQ 320
Qy 354 YHFVANDITASFTSKEDMKFNFTYHCLNEEIKAEI-----EKKYAK-VNSTHXYGD 406
Db 321 NYDKSNHMTIPTNEDEISKIISVETKMDKDEILSKYNTVIDFNKKYKETYNSHESQFTE 380
Qy 407 LKYPEKTDGGLYLVWQPLIQNRLLDAKNKLANNETYSRRSRQAESTTD--PMWEMTNGAG 464

381 L-----TD-----KIKAEVSDRELKKCEQSFNDNKSILNETKNSIE 416
Qy 465 GEYSSENSITVAQVQAYDNLIRI--NNILEDLSKAWCREHQRAALVWNELSKINPTSV 523
Db 417 KEYQNINTLKKVD-EY-----IKVCKSTKESITKFSSKQTLKMDLNLQNIKVTKETNSI 469
Qy 524 SMYLRPVSAKRIGDVISVSNCIWVDTSVSLHKSRLLSASDEKCFSPRPVTFKWNDS 583
Db 470 DKYIEKFEQILTKQTKLENKF-----TEFSLNH-----EANNNE-----LIKFS 513
Qy 584 TIYKQGLGVNNEILLTTTLETQENTYFQAKTDMYIKYVYEHKLTVPSSII--TTLD 641
Db 514 --LKANLGINEENLNYQFTE--KEKT---FNDIKEKNHINEISIKIEIKHASYNIS 566
Qy 636 SITTLDTFIALNFTLLENVDFKVIETRDEKRLSNVDFIETMFREYNYAQRVSGRLK 695
Db 567 EETERE--IGIN--IESLTKVFE--KVKNVTNLNKIKEKLKHY-----D 606
Qy 696 LLDLSTNRN-QFVDAFGSLMDDLGAAGVGTNNVAVSGVATL-----PSSIVTGFNFIKN 748
Db 607 FSDFGKEGNIKYTDKIKKINDDIWVSSQIDQHINGLDDIQKSESIVSEMKGQINKLEK 666
Qy 749 PFGGMLMIIVVIGLFAIYFLTKTKIYETA--PIKMIYPEIDKLKEREKGSBIAPISEE 806
Db 667 VSNTAISNDNVEGI-----KKQOIIVTKIDKKNYIEINKLSEISKIEKONTSL 719
Qy 807 ELERIVLAMIHOONSHMETKTRDKPKDSILTRQANMLKRSY----SNL-KNAESVE 860
Db 720 KVKDINLSYQNLGNLFLEQIDBEKKK-----AENTIKSMEAYIDDLNKKKSQIE 772

RESULT 43
F28885
hypothetical protein UU482 [imported] - Ureaplasma urealyticum
C:Species: Ureaplasma urealyticum
C>Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 20-Aug-2000
C:Accession: F28885
R:Glass, J.I.; Lefkowitz, E.J.; Glass, J.S.; Heiner, C.R.; Chen, E.Y.; Cassell, G.H.
submitted to GenBank, February 2000
A:Description: The complete sequence of Ureaplasma urealyticum: Alternate views of a min
A:Reference number: A82870
A:Accession: F28885
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-4688 <GLA>
A:Cross-references: GB:AE002145; GB:AF222894; NID:G6899476; PIDN:AAF30894.1; GSPDB:GN0012
A:Experimental source: serovar 3; biovar 1
A:Genetic code: SGC3

Query Match 3.2%; Score 146; DB 2; Length 4688;
Best Local Similarity 19.3%; Pred. No. 10;
Matches 184; Conservative 120; Mismatches 349; Indels 300; Gaps 42;
Qy 92 GVGDFRFTQDTHVCDDSDMVHSE-----GILLIYKQNIIPFMRVYKRVKVT--TSTV 144
Db 2337 GVGNIETIQDRVHNLSAKIRFELDLNVLSDNQATITNNNTTSKAVITDQNKY 2396
Qy 145 YNGIYSDSITNOHTFYKSIET---PWETEK-----MDTIYOCFN----- 179
Db 2397 LEATFNLVLNKDTIINKIEFNTPKKNVASKNGINTNVIYDATNLINNDLKITGLPHT 2456
Qy 180 -----SURLNTGGNL-----LTYVDRDDIN--MTVFLQPVQDVTDPVKRYGS 220
Db 2457 LKBEFANNKNTISVSLDNTNNHISKNLVFIKAFDSNDGQSVLTNPAAADKIVTNNNNK 2516
Qy 221 PEL-----YLEPGWFGWGSYRRRTTVNCELMDMFARSNPPDFFTATGTVEM 268
Db 2517 KELTFNLTLNLSNRQYIFKGLYVYVSSNQTNID--ENNKFEKNSNVYKITVKPTITTIQ 2575
Qy 269 SPWSCEDDHENKMKHKEKPFVSVNNYKVVYQNRGTVPGLKTRIFLDEEYTLSEWHL 328

Db 2576 NGNWTFPQNAQK-----FKFNINSNDVDF-----STDLDATITSDQHTKKTITTKL 2625
Qy 329 K---NMSYCPITLWKAIFYNG-----IQTEHSGSHFVANDITASF 365
Db 2626 KQKDNQWYEDTISLDLAYNDTYKLDISITKPNWAFANLKIQIE-----NKEQISF 2676
Qy 366 TTSKEDMKFNTTYHCL--NEEKABIEKKY-AKVNSTHSGYKGLKYFKYKTDGGLYLVWQP 422
Db 2677 TTQSGTQLVSIITSHLWMDQPNASQQTITAKVSGVNDLY----- 2718
Qy 423 LIONRLDANKLNNETYGRSRQRAESTTDPMMWTGAGGEYSSSENSITVAQVQYAY 482
Db 2719 --NNR-----KILVVEYRSNNOKSLVESN-----ELTQKQDQOYIFTLPISVANRQYSF 2767
Qy 483 DNLIRIRINILEDLSKAWCREQHRAALVMNELSKINPTSVMSIYNRPVSAKRIGDVISV 542
Db 2768 KEIKITSN-----NNFETLNNSTNVTSFN--VNASKTOIVDN 2805
Qy 543 SNICVVDQTSVLHSLRLLSASDEKCFRPPVTFKFMNDSTIYKGLGVNNEILLTTY 602
Db 2806 TN-LATNITSTATITTYKLKSDHVFQVGSIIITYLKSNDDE-----RQEISYTKTI 2856
Qy 603 LETQCENTEYFOAKTDMYIK--NYEHLKT-----VPLSSITTL-----DTFIAL 646
Db 2857 TSISDGEATVSFNTDQTLKEEANYKLIKVGFRKPTLAYTNINNDANNVIFEDNNSY 2916
Qy 647 NF-TLLENVDKVIELYTRD-----EKRLSNVF-----D 674
Db 2917 NFKTLI--VDHKVTNVSSNDSINTTTQTVNIDIDGQRTWINKKIQLVYTSNDGEILLSD 2974
Qy 675 IETMFREYNYAQRVSGLRKDLDLSTNRNQFVDAFGSLMDD-----LGAVGQTV 724
Db 2975 QXTLLWANNHYSPFLSNLKH-----NRKYLKEVRIINDNKTSIIFHLKNGADWFI 3027
Qy 725 VNAGSVATLFGSIVTGF-----INFIKNP-----FGMLMI--- 756
Db 3028 VNKTQIS--ISGIIPIARAKNNLQSTQIRFILNDPDNVLSNEEAEIYNGENLSVRAK 3085
Qy 757 IVVIG-----VLFAYFLPKTKIYETAPI-----KMIY--PEIDKUK 792
Db 3086 KVIQSGKYLETFNNLALNQDTIINSITFINKPIKAATNIGIDNSKNKYNTNINPLK 3145
Qy 793 EREGKEIAPISEEELERIVLAWHIHOONSHMETKTRDKPK--DSLTRAQNM 843
Db 3146 IDNFSVIGVQSDTNKRITL-----EYNNKTNVKNLEKANDNL---QNL 3190
RESULT 44
A38539
p101 protein precursor - Mycoplasma hyorhinis
C:Species: Mycoplasma hyorhinis
C>Date: 31-Jul-1991 #sequence_revision 31-Jul-1991 #text_change 07-Dec-1999
C:Accession: A38539
R:Yogev, D.; Watson-McKown, R.; McIntosh, M.A.; Wise, K.S.
J. Bacteriol. 173, 2035-2044, 1991
A:Title: Sequence and TnphoA analysis of a Mycoplasma hyorhinis protein with membrane ex
A:Reference number: A38539; MUID:91161521; PMID:1848219
A:Accession: A38539
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-888 <YOG>
A:Cross-references: GB:M60787; NID:g150162; PIDN:AAA25422.1; PID:g150163
C:Genetics:
A:Genetic code: SGC3
Query Match 3.2%; Score 145.5; DB 2; Length 888;
Best Local Similarity 19.4%; Pred. No. 0.93;
Matches 163; Conservative 114; Mismatches 293; Indels 269; Gaps 38;
Qy 141 TSTVY--NGIYSDSINQHTFYKSIPEWETEKMDTIYQCFSRLRLMTGNL----- 189
Db 110 TRWVLLSGLYVDN--QAATLIKSPF-EVNLKNIIVKA-----TNAGNQLIFNPMLSAN 161

Qy 190 ---LTYVDRDDINNTVFLQPV--DGV--TPDVKRYGSOPELYLEPGFWGVSRRRTTV-- 240
Db 162 EKLKLTFFKDKNAKETTPEASTNQDGVNINSEKINLO-----GNYQLVASVIA 210
Qy 241 -----NCELMDM--FARSNPPDFPFVATGD-----TVEMSPFWSGE 275
Db 211 KNNYSLDFNSQVIEWSFPFKSNP-----VPTSDVMIKNDKVAINKINTLELS----- 259
Qy 276 DDHENKMEK--PWFVSVINNYKV--DYQNGRTVPLGKTRIFLDR-----EYTLSE 325
Db 260 -DHLKFPANIKDPSQALDYEVWLTAKNQKPAELLASTAKIMDKLVFTFDNLTSNTS 318
Qy 326 KHLKMSYCPITLWKAIFYNGIQTEHSGSHFVANDITASF TTSKEDMKFNTTYHCLN-- 383
Db 319 YQISNISY-----KNDYTDKILLDPSINKTFTTNSKPLDFD 354
Qy 384 -BEIKAEIEKKYAKVNSTHSGYKGLKYFKTDGGLYLVWQPLIQNRLLDKANKLN----- 436
Db 355 PSNFALKIKDNHYKVNKLSS-----KONLEANOAVNVVFTSE 393
Qy 437 NETYSRRSRQRAESTTDPMMWTGAGGEYSSSENSITVAQVQYAYD-----NLRIRINI 492
Db 394 DKTLHPAKITQDSTND-----YKIELELDNLQDDQIYSLEMIKALKPTKA 440
Qy 493 LEDLSKAWCREQHRAALVMNELSKINPTSVMSIYNRPVSAKRIGDVISVNCIVVDQTS 552
Db 441 LKNLGAQDNVIYNAKDDFKISFIN-EKWLNIITLTREVDPKTLTD----- 485
Qy 553 VSLHSLRLLSASDEKCFRPPVTFKFMNDSTIYKGLGVNNEILLTTYLETQCENTEY 612
Db 486 --KHSSYKI-----KLNFSK--IPTYLNHKIALEYQ-SKND-LATTSEVEIKDSSLTD 534
Qy 613 YFOAKTDMYIYKVEHLKTVPLSSITTLDTFIALNFTLLENVDFKVIELYTRDEKRLSNV 672
Db 535 YTLDLNNLVNKEYE-----LKFLIKDNTNNKVIKITKD-LNLNDS 575
Qy 673 FDIETMFREYNYAQRVSGLRKDLDLSTNRNQFVDAFGSLMDDLGAVGQTVNNAVSGVA 732
Db 576 LKILQTFNFOIKERTINSLNAS--VDINTDLNNSKVDF--SONKQVKVYQEVNNSN--- 630
Qy 733 TLFSSIVTGFNFIFKPFPGGMLMIIVIGLVFAIVFLTKTKI-----YETAP 780
Db 631 ---THQITGTTKDNNNNNKQL-----DLTFNLELNKKYIKTLEVTNSANDAHNVIV 681
Qy 781 IKMIYPEIDKUKEREGKSEIAPISEEELERIVLAWHIH-----QONSHME 825
Db 682 FNLFTP-----WELQNOFEVSSISDLNLKTKQLQDFQINTNIDLTNSKYKLMKNNQNV 736
Qy 826 TKTRKDKPKDSLTRAQ-----NMLKRSYGNLKNKAESEVEMLN 863
Db 737 KEIVTNQKDSITFNKQNTNIISVKLSSIFNQEGKYLIKLLIANNNNNNFTNAKEISUSN 795
RESULT 45
B42771
reticulocyte-binding protein 2 - Plasmodium vivax (fragment)
C:Species: Plasmodium vivax
C>Date: 12-Mar-1993 #sequence_revision 12-Mar-1993 #text_change 09-Sep-1997
C:Accession: B42771
R:Galinski, M.R.; Medina, C.C.; Ingravallo, P.; Barnwell, J.W.
Cell 69, 1213-1225, 1992
A:Title: A reticulocyte-binding protein complex of Plasmodium vivax merozoites.
A:Reference number: A42771; MUID:92315338; PMID:1617731
A:Accession: B42771
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-1252 <GAL>
A:Cross-references: GB:M80898; NID:g160627; PID:g160628
A:Experimental source: strain Belem, merozoites
C:Genetics:
A:Gene: RBP2

```
Query Match      3.2%; Score 144; DB 2; Length 1252;
Best Local Similarity 19.0%; Pred. No. 1.9;
Matches 165; Conservative 138; Mismatches 319; Indels 246; Gaps 42;

Qy 109 SDMVHSEGLIYKONI-IPMFV-----RKYRKVT-----TSTV--YNGIYSDSITN 155
   ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||:
Db 506 STITNIEGALKESGNGVEIGFLEKLEIGKRNKLVDTTKSINSTVGNFSSLFNNFDLN 565

Qy 156 QHTFYKSIEPWETEKMDTIYOCF-----NSLRNTCGNLLTYVDRDDIN 199
   ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||:
Db 566 QYDFNKNINDYE-NKMGIEIYNEFEGSLNKISENLNASENTSDYNSAKTLRLLEAQEKVN 624

Qy 200 MTFVLPQPDVGTDPVKRYGS-----OPELYLEPGWFGSYRRRTTV 240
   ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||:
Db 625 LNKKEEANKYLRDVKKVESFRFIPNWKESLDKINEMIKKEQLTVNEG--HGNVKQLVEN 682

Qy 241 NCELMDFARSNPPFFVTATGDTVEMSPFWSGEDDHEN-KMHEKPFVSVINNYKVVD 299
   ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||:
Db 683 IKELVD---ENNLSDLKQATGKNEEQKI-----THSTLKNKAKTILGHVDTSAKYVG 733

Qy 300 YONRG-----TVPLGKTRI-----FLDREYTLSEKHLKMSYCPPLTWKAFYNGIOT 348
   ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||:
Db 734 IKITPELALTELLGDALKTAQELKFESKNNVLETEMNSKNTNELD-----HKNIQD 787

Qy 349 EHSGSYHFVANDITASFTTSKEDMKEF-----NTTY--HCLNE-----BIKAEIEKKA 395
   ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||:
Db 788 AYKVALEILAH--SDEIDTKOKOSKLEMGNOIYLVVNLINQYKNKISSIKKEEAVSV 845

Qy 396 KVNSTHSGYGLDYF-----KTDGGLYLVMOPLIQ---NRLLDARNKLNNETYSRRSRQ 447
   ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||:
Db 846 KIGNVSKHSELSKITCSKSYDNIIALEKOTELQNLNRSFTQKTNSTNSDSKLEKIKTD 905

Qy 448 AESTTDPMENTGAGAGEYSENSITVAQVOYAYDNLIRINNILEDLKAWCREOHRA 507
   ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||:
Db 906 FESLNALKTLEGEVNAKASSDN-----HEH-- 932

Qy 508 ALVWNELSKINPTSVMSMIYNRPVSAKRIGDIVISVNCIVVDQTSV-SLHKSLRLLSASD 566
   ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||:
Db 933 --VOSKSEPNP-----ALSE-IEKEETDIDSLNTALDELKKG 968

Qy 567 EKC-FSRPPVTFKPMNDSTIYKGOLGVNNEILLTTYLETQENTYFYFOAKTDMYIYKN 625
   ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||:
Db 969 RTCEVSR---YKLIKOT-----VTKEISDDTELINTIEKNVAYL-----AYIKKN 1011

Qy 626 Y-----EHLKVPLS-----SITLDTFIALNFTLLENVDFKVIELY 662
   ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||:
Db 1012 YEDTVQDVLTLNEHFNTKQVSNHEPTNPKSNKSEBELTKAVTDSKTIISKLGKGVIEVN 1071

Qy 663 TRDEKRL--SNVFDIETMFREYNYAQRVGLRK-----DLDDLSTNRNQFVDA-- 709
   ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||:
Db 1072 ENTWNTIESSAKEIEAL---YNELKNKKTSLNEIYQTSNEVKLQEMKSNADKYIDVSKI 1128

Qy 710 FGSLMDDLGAQGVTVNAVSGVATLFGSSIVTG--FINPIKNPFGGMLMIIVIGVLPFAIY 767
   ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||:
Db 1129 FNTVLD-----TQKSNIVTNQHSINNVDKLGKQLQELIDADSSFTLE 1171

Qy 768 FLTKKTIYETAPIKMIYPEIDKLEREGKSEIAPISEEELERIVLAWHIHQONSHMETK 827
   ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||:
Db 1172 SIKKFNEIY--SHIKTNIQEQL-QQTNKSEHDNVAKH-EKIV---HLINRVESLKG 1224

Qy 828 TRKDPKDSILTRAQNMLRKRSYGNLKN 855
   ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||:
Db 1225 VKNHDDDOYMKKLNASLN-----DNKN 1248
```

Search completed: October 28, 2003, 15:31:01
Job time : 59 secs

GenCore version 5.1.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 28, 2003, 15:27:11 ; Search time 26 Seconds
(without alignments)
1564.542 Million cell updates/sec

Title: US-10-055-364-24

Perfect score: 4547

Sequence: 1 MAGSLKRGSLALWLYQV.....KRGYSNLKNAESVEMLNTL 865

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_41.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	1949.5	42.9	808	1 VGLB_HSVSA	P24905 herpesvirus
2	1848.5	40.7	857	1 VGLB_BEV	P03188 epstein-bar
3	1134	24.9	854	1 VGLB_RHCM6	P89053 rhesus cyto
4	1128	24.8	907	1 VGLB_RCMVT	P13201 human cytom
5	1121.5	24.7	906	1 VGLB_RCMVA	P06473 human cytom
6	1084.5	23.9	822	1 VGLB_HSV7J	P52352 human herpe
7	1055.5	23.2	830	1 VGLB_HSV6U	P28864 human herpe
8	1050.5	23.1	830	1 VGLB_HSV6G	P36319 human herpe
9	1040.5	22.9	830	1 VGLB_HSV6Z	P36320 human herpe
10	1034	22.7	928	1 VGLB_MCMVS	P27171 murine cyto
11	1008	22.2	944	1 VGLB_HSVT2	Q9wr15 herpesvirus
12	999	22.0	901	1 VGLB_GPCMV	Q69024 guinea pig
13	900	19.8	920	1 VGLB_HSVSM	Q04463 herpesvirus
14	876.5	19.3	865	1 VGLB_HSVMD	P18538 marek's dis
15	857	18.8	932	1 VGLB_HSVBC	P12640 bovine herp
16	849.5	18.7	904	1 VGLB_HSV1K	P06437 herpes simp
17	845.5	18.6	903	1 VGLB_HSV1F	P06436 herpes simp
18	844	18.6	904	1 VGLB_HSV2H	P08666 herpes simp
19	841	18.5	904	1 VGLB_HSV11	P10211 herpes simp
20	839	18.5	904	1 VGLB_HSV23	P08663 herpes simp
21	836.5	18.4	904	1 VGLB_HSV1P	P08665 herpes simp
22	836	18.4	883	1 VGLB_ILTVS	P27415 infectious
23	835	18.4	980	1 VGLB_HSVEB	P28922 equine herp
24	833	18.3	885	1 VGLB_HSV2S	P24994 herpes simp
25	832	18.3	868	1 VGLB_VZVD	P09257 varicella-z
26	830	18.3	883	1 VGLB_ILTVT	P24904 infectious
27	829	18.2	873	1 VGLB_ILTV6	Q02409 infectious
28	824	18.1	980	1 VGLB_HSV1	P18050 equine herp
29	824	18.1	980	1 VGLB_HSV1A	P18551 equine herp
30	821	18.1	933	1 VGLB_HSV1	Q04463 herpesvirus
31	819.5	18.0	919	1 VGLB_HSV1E4	P17472 equine herp
32	805	17.7	917	1 VGLB_HSVB2	P12641 bovine herp
33	777.5	17.1	979	1 VGLB_HSV1	P25218 equine herp

RESULT 1

ID	VGLB_HSVSA	STANDARD;	PRT;	808 AA.
AC	P24905;			
DT	01-MAR-1992 (Rel. 21, Created)			
DT	01-MAR-1992 (Rel. 21, Last sequence update)			
DT	16-OCT-2001 (Rel. 40, Last annotation update)			
DE	Glycoprotein B precursor.			
GN	8 OR KCRF1.			
OS	Herpesvirus saimiri (strain 11).			
OC	Viruses; dsDNA viruses, no RNA stage; Herpesviridae;			
OC	Gammaherpesvirinae; Rhadinovirus.			
OX	NCBI_TaxID=10383;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=90163221; PubMed=2154888;			
RA	Albrecht J.-C., Fleckenstein B.;			
RA	"Structural organization of the conserved gene block of Herpesvirus saimiri coding for DNA polymerase, glycoprotein B, and major DNA binding protein.";			
RT	binding protein.";			
RL	Virology 174:533-542(1990).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=92333688; PubMed=1321287;			
RA	Albrecht J.-C., Nicholas J., Biller D., Cameron K.R., Biesinger B., Newman C., Wittmann S., Craxton M.A., Coleman H., Fleckenstein B., Honess R.W.;			
RA	"Primary structure of the herpesvirus saimiri genome.";			
RT	J. Virol. 66:5047-5058(1992).			
RL	!- SIMILARITY: BELONGS TO THE HERPESVIRUSES GLYCOPROTEIN B FAMILY.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).			
CC	-----			
CC	EMBL; X64346; CAA45631.1; --			
DR	EMBL; M31122; AAA46164.1; --			
DR	InterPro; IPR000234; Glycoprot_B.			
DR	Pfam; PF00606; Glycoprotein B; 1.			
DR	ProDom; PD000693; Glycoprot_B; 1.			
KW	Glycoprotein; Transmembrane; Signal.			
FT	SIGNAL 1 17 POTENTIAL.			
FT	CHAIN 18 808 GLYCOPROTEIN B. (POTENTIAL).			
FT	DOMAIN 18 657 EXTRACELLULAR (POTENTIAL).			
FT	TRANSMEM 658 673 POTENTIAL.			
FT	TRANSMEM 679 698 POTENTIAL.			
FT	TRANSMEM 702 722 POTENTIAL.			
FT	DOMAIN 723 808 CYTOPLASMIC (POTENTIAL).			
FT	CARBOHYD 30 30 N-LINKED (GLCNAC. . .) (POTENTIAL).			
FT	CARBOHYD 158 158 N-LINKED (GLCNAC. . .) (POTENTIAL).			
FT	CARBOHYD 239 239 N-LINKED (GLCNAC. . .) (POTENTIAL).			
FT	CARBOHYD 251 251 N-LINKED (GLCNAC. . .) (POTENTIAL).			

P08355 pseudorale
P17471 bovine herp
Q00799 plasmodium
P28742 saccharomyc
P24339 schizosacch
P47437 mycoplasma
P36022 saccharomyc
Q02455 saccharomyc
Q23639 caenorhabdi
P19321 clostridium
Q27171 paramedim
Q04956 plasmodium

ALIGNMENTS

```
FT CARBOHYD 285 285 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 331 331 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 344 344 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 355 355 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 361 361 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 332 332 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 569 569 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 587 587 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 598 598 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 808 AA; 91694 MW; FA1BCB89B9D5672E CRC64;

Query Match 42.9%; Score 1949.5; DB 1; Length 808;
Best Local Similarity 46.6%; Pred. No. 3.7e-112;
Matches 373; Conservative 159; Mismatches 236; Indels 33; Gaps 10;

Qy 71 EONKNIYGSPTFYRVCSAGSGVDVFRFQTDHVCPCDASDMVHSEGLLIYKQNIIPFMF 130
Db EKNKTQAIYQBYFYRVCSASTTGELFRFLDRDTCPTSEDKVHKEGILLVYKKNIVPYIF 87

Qy 131 RVRYKRVVITSTVYNGIYSD--SITNQHTFYKSIETWETEKMDTIYOCFNSLRNLNCGN 188
Db KVRKYKITTYSVRFNGWTRREGVAITNKLWSRAVPKYIEIDMDKTYOCHNCMOIEVNGM 147

Qy 189 LLTVVDRDDNMTVFLQPDGVTPDVKRYGSQPELYLEPGWFGSYRRRTTVNCELMDMF 248
Db LNSIYDRDGNKNTVDLKFVGLTGATRYISQPKVPADPGWLGTYTRITVACEIVDMF 207

Qy 249 ARSNPPDFVATGDTVEMSPFSGEDDHENKHEKPFVSVINNYKVVDYQNRGTVP 308
Db ARSADPTYFTVLTGDTVEVSPFCVDNCPNATDVLVS--VOIDLNHTVVDYGNRATSSQ 265

Qy 309 GKTRIFLDREYTLSEKHLKMSYCPCLTWKAFYNGIOETHSGSYHFVANDITASTTS 368
Db HKKIFAFHTLDYSVSEAVNKSASVCSVMFQKSFQRAIQTEHDLTYHFIANETAGSTV 325

Qy 369 KEDMKFNTVYHCLNEIEKAEIKKYAKVNSTSKYGDLYKFDGGLYLWVQPLIQNL 428
Db KEPLANTSNDYCNLMTHTNTLEDKIAVNTHTPNGTAYYQTEGGMLVWQPLIAEL 385

Qy 429 LDANKLNNE-----TYRRSRROAESTTDPMMETNGAGGEYSSENSITVAQVQYA 481
Db EAMLEATTSPVTPSAFTSSSRKRAIRSIRDV-----SAG-----SENNVFLSQIQYA 434

Qy 482 YDNLRIINNILEDLSKAWCEQRAALVWNLKSNPTSVMSMIYNRPVSAKRIGDVIS 541
Db YDKURQSIINNVLLELATWCREQVQRTVMVYEIAKINPTSVMTAIYKPVSRKALGDVIS 494

Qy 542 VSNICVVDQTSVLSHLKSLRLLSASDEKCFRPPVTFKFMNDSTIYKQGLGVNNEILLTT 601
Db VTECINVDQSSVSIHKSILK--TENNDICYSRPPVTFKFNSSQLFKQLGARNEILLS 552

Qy 602 YLETCQNTYFYFOAKTDMYIKNYEHLKTVPLSSITLDTFIALNFTLENVDKVEL 661
Db LVENCHQNAETFTAKNETYHFNKYHVHETLPVNNISTLDTLALNLTFTENIDFAVEL 612

Qy 662 YTRDEKLSNVEDLETFRFENYVYQVSGLRKDLDLDS--TNRNQVDAFGSLMDLGA 719
Db YSSGERKLANVFDLETFRFENYVYQVSGLRKD--FDNSQRNDRIRIQDFSEILADLGS 671

Qy 720 VQOTVNAVSGVATLFSISVTGFINFKNPFPGGMLMIVIGVLFAIYFLTKTKIYETA 779
Db IGKIVNVASGAPSLFGVITGILNFKNPLGNGFTFELLGAVILLVILLVRRTNNSQA 731

Qy 780 PIRMIYPEIDKLKREKGSIAPISEBELERIVLAMHIHQONSHMETKTRKDPKDSILTR 839
Db PIRMIYDPVEK-----SKSTVTPMETIKQILLGMHNMQOEAYKKKEQRAARPSIFRQ 786

Qy 840 -AQNMLKRGSGYNLKAESV 859
Db AAEFTFLRRSGYKQISTEDKI 807
```

RESULT 2

```
VGLB_EBV
ID VGLB_EBV STANDARD; PRT; 857 AA.
AC P03188;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Glycoprotein GP110 precursor (GP115).
GN BALF4.
OS Epstein-Barr virus (strain B95-8) (Human herpesvirus 4).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Gammaherpesvirinae; Lymphocryptovirus.
OX NCBI_TaxID=10377;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=85035713; PubMed=6092825;
RA Bankier A.T., Deininger P.L., Farrell P.J., Barrell B.G.;
RT "Sequence analysis of the 17,166 base-pair EcoRI fragment C of B95-8
RT Epstein-Barr virus.";
RL Mol. Biol. Med. 1:21-45 (1983).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=84270667; PubMed=6087149;
RA Baer R., Bankier A.T., Biggin M.D., Deininger P.L., Farrell P.J.,
RA Gibson T.J., Hatfull G., Hudson G.S., Satchwell S.C., Seguin C.,
RA Tuffnell P.S., Barrell B.G.;
RT "DNA sequence and expression of the B95-8 Epstein-Barr virus genome.";
RL Nature 310:207-211 (1984).
RN [3]
RP IDENTIFICATION OF PROTEIN.
RX MEDLINE=87112946; PubMed=3027378;
RA Gong M., Ooka T., Matsuo T., Kieff E.;
RT "Epstein-Barr virus glycoprotein homologous to herpes simplex virus
RT gB.";
RL J. Virol. 61:499-508 (1987).
CC -I- SIMILARITY: BELONGS TO THE HERPESVIRUSES GLYCOPROTEIN B FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; V01555; CAA24806.1; -
DR PIR; A03749; QOBEIL.
DR InterPro; IPR000234; Glycoprot B.
DR Pfam; PF00606; Glycoprotein_B; 1.
DR ProDom; PD000693; Glycoprot_B; 1.
KW Glycoprotein; Transmembrane; Signal; Late protein.
FT SIGNAL 1 22 POTENTIAL.
FT CHAIN 23 857 GLYCOPROTEIN GP110.
FT DOMAIN 23 688 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 689 704 POTENTIAL.
FT TRANSMEM 710 729 POTENTIAL.
FT TRANSMEM 733 753 POTENTIAL.
FT DOMAIN 754 857 CYTOPLASMIC (POTENTIAL).
FT CARBOHYD 163 163 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 290 290 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 329 329 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 348 348 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 395 395 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 436 436 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 563 563 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 629 629 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 857 AA; 95639 MW; D9BCE9487D8A1411 CRC64;

Query Match 40.7%; Score 1848.5; DB 1; Length 857;
Best Local Similarity 46.1%; Pred. No. 6.4e-106;
Matches 368; Conservative 148; Mismatches 241; Indels 41; Gaps 12;

Qy 81 STFPYRVCSAGSGVDVFRFQTDHVCPCDASDMVHSEGLLIYKQNIIPFMFRVRYKRV 139
```


QY 419 VMOPLIONRLLDKAKN---KLNNETYSRRSRQAEISTTDPMMEMTNGAGGEYSSE--NSI 473
 Db 403 FWLPVKSEKSIEMEEELAVANNYNTSSRRKRSTDSASD-----SNKTSSEVLKSI 452
 QY 474 TVAQOVAYONLRIIRINILEDLSKAWCREQRAALVWNLKSKINPTSVMSMIYNNRPVSA 533
 Db 453 VYAOQVYTYTLRINYINRALRQIAEAWCKQKRTLEVFVKLSKINPMSAMLSAIYDKPIAA 512
 QY 534 KRIGDVISVNCIIVDQTSVSLHKSLLSASDEKFCSPRPVTFKFWNDSTIYKQQLGVN 593
 Db 513 RFVGDVISLAKCDEVQNSVKVLRDMR--TKESGVCYCRPVVLYTFKNSHVOYQQLGEY 570
 QY 594 NEILLTYYLETQENTYFYFOAKTDMYIYKNYEHKTVPLSSITLTDPIALNFILLEN 653
 Db 571 NEILLGHRHTEACEYSLKIKIYAGNSSYEVVDVLYKRMIFLDSISTVDTMISLDIDPLEN 630
 QY 654 VDFKVIELYTRDEKRLSNVFDIETMPREYNVYAQRVSLGRKDLDLSTNRNQFVDAPGSL 713
 Db 631 TDFKALELYSEDELRSNVFLEDDIMEFNTYKQRMITHVEGKVPDVPKG---YLRGLDDM 687
 QY 714 MDDLGAQGVTVNNAVSGVATLFSISVTGTFNFKNPFEG---MLMIIVIGVLPFAIFYLT 770
 Db 688 MSGLSAGKALGVAIGAAGAVASFEVGVGFIKNPFGSFTVILFLAVLGVLYIYMRQ 747
 QY 771 KTKIYETAPIKMIYP-----EIDKLERECKSEIAPISSEELERIV 812
 Db 748 KRA--YB-KPHEHFPYVVPPTVKEAPPYSQYENIKEK-AASATKFSLEEAQOML 803
 QY 813 LAMHIHQONGHMETKTRKDPKD-----SILTRAQNMRLKRGSGYNLKNAESV 859
 Db 804 LA--LQKLDQEKRAADEDFASNGOSAGFLDLRLN--RRRGYQKIQNEIYEV 854

RESULT 4
 VGLB_HCMVT
 ID_VGLB_HCMVT STANDARD; PRT; 907 AA.
 AC P13201;
 DT 01-JAN-1990 (Rel. 13, Created)
 DT 01-JAN-1990 (Rel. 13, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Glycoprotein B precursor [contains: Glycoprotein GP55].
 GN GB OR UL55.
 OS Human cytomegalovirus (strain Towne).
 OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
 OC Betaherpesvirinae; Cytomegalovirus.
 OX NCBI_TaxID=10363;
 RN [1]
 RP MEDLINE=89045645; PubMed=2460994;
 RX Spaete R.R., Thayer R.M., Probert W.S., Masiarz F.R.,
 RA Chamberlain S.H., Rasmussen L., Merigan T.C., Pacht C.;
 RT "Human cytomegalovirus strain Towne glycoprotein B is processed by
 proteolytic cleavage."
 RL Virology 167:207-225(1988).
 CC -!- SIMILARITY: BELONGS TO THE HERPESVIRUSES GLYCOPROTEIN B FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; M22343; AAA45920.1; -;
 CC PIR; A31288; VGBETE.
 CC InterPro; IPR000234; Glycoprot_B.
 CC Pfam; PF00606; Glycoprotein_B_1.
 CC ProDom; PD000693; Glycoprot_B_1.
 CC Glycoprotein; Transmembrane; Signal.
 CC SIGNAL 1 24 POTENTIAL.
 FT CHAIN 25 907 GLYCOPROTEIN B.

FT CHAIN 461 907 GLYCOPROTEIN GP55.
 FT DOMAIN 25 707 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 708 723 POTENTIAL.
 FT TRANSMEM 729 748 POTENTIAL.
 FT TRANSMEM 752 772 POTENTIAL.
 FT DOMAIN 773 907 CYTOPLASMIC (POTENTIAL).
 FT CARBOHYD 68 68 N-LINKED (GLCNAC..) (POTENTIAL).
 FT CARBOHYD 73 73 N-LINKED (GLCNAC..) (POTENTIAL).
 FT CARBOHYD 85 85 N-LINKED (GLCNAC..) (POTENTIAL).
 FT CARBOHYD 208 208 N-LINKED (GLCNAC..) (POTENTIAL).
 FT CARBOHYD 281 281 N-LINKED (GLCNAC..) (POTENTIAL).
 FT CARBOHYD 286 286 N-LINKED (GLCNAC..) (POTENTIAL).
 FT CARBOHYD 302 302 N-LINKED (GLCNAC..) (POTENTIAL).
 FT CARBOHYD 341 341 N-LINKED (GLCNAC..) (POTENTIAL).
 FT CARBOHYD 383 383 N-LINKED (GLCNAC..) (POTENTIAL).
 FT CARBOHYD 405 405 N-LINKED (GLCNAC..) (POTENTIAL).
 FT CARBOHYD 409 409 N-LINKED (GLCNAC..) (POTENTIAL).
 FT CARBOHYD 417 417 N-LINKED (GLCNAC..) (POTENTIAL).
 FT CARBOHYD 447 447 N-LINKED (GLCNAC..) (POTENTIAL).
 FT CARBOHYD 452 452 N-LINKED (GLCNAC..) (POTENTIAL).
 FT CARBOHYD 456 456 N-LINKED (GLCNAC..) (POTENTIAL).
 FT CARBOHYD 466 466 N-LINKED (GLCNAC..) (POTENTIAL).
 FT CARBOHYD 555 555 N-LINKED (GLCNAC..) (POTENTIAL).
 FT CARBOHYD 586 586 N-LINKED (GLCNAC..) (POTENTIAL).
 SQ SEQUENCE 907 AA; 101953 MW; E6F07B7742D359A2 CRC64;
 Query Match 24.8%; Score 1128; DB 1; Length 907;
 Best Local Similarity 30.9%; Pred. No. 1.2e-61;
 Matches 286; Conservative 184; Mismatches 353; Indels 102; Gaps 25;
 QY 14 LMYVOALVYLSIAETG--VTSPNTATWSTESPLTGHVGHDS--HGERNNENRDS 69
 Db 5 IWCL--VVCNLCIVCLGAASVSSSTRGTSATHS-----HSSHTTAAHRSRGSVQRVT 58
 QY 70 EEQ-----NKNIYG-----SPSTFPYVCSASGVGVDFRQTDHVC---PDA 108
 Db 59 SSQTGVSHGVNETIYNTLLKGVGVGVTTKYPRVCSMAQGTDLIRFERNIVCTSMKPIN 118
 QY 109 SDMVHSEGLILLYKQNIIFPMFRVRYKRVKVTSTVYNGIYSDISITNQHTFYKSIQWET 168
 Db 119 EDL--DEGIMVYKRNIVATFKRVYQKLTFRRSYAVIHTTYLLGNSNTEYVAPPWMEI 176
 QY 169 EKMDTIYQCFNSLRNLNTGNLLTYVDRDDI--NMTVFLQPVGVTPDVKRYGSPELYLEP 227
 Db 177 HHINSHSQCYSSYSRVIAGTVFVAYHRDSEYENKTMQMPDDYNSHTSTRIVTVKQDQHSR 236
 QY 228 GMFWGSYRRRTTVNCELMDMFARNSPPFFFTATGDTVEMSPFSGEDDHENKMEKPW 287
 Db 237 GSTW--LYRETCNLNCWVTITTARSKYPYHFFATSTGDVVDISPFYNGTNRNASYFGENAD 295
 QY 288 FVSVINNYKVVDYQNRGTVPGLKTRI--FLDREYTLSE--KULKMSYCPPLTLWKAFYN 344
 Db 296 KFFIFPNYTIYDFGRPNSALETHRLVAFLEADSVISMDIQDEKNVT--COLTFWEASER 354
 QY 345 GIOTHSRGYVHFVANDITASFSTTSKEDMKEFNTTYHCLNEEKABIEKY--AKVNSTHSHK 403
 Db 355 TIRSEADSYHSSAKMTATFLSKQEVNMSDALDCVRDEAINKLQQLFNTSYNQTYEK 414
 QY 404 YGDLKYFTDGGLYLVWQPLIQNRLLDAKNLNNETYSRRSRQAEISTTDPMMEMTNGA 463
 Db 415 YGNVSVFETTGGLVWFVWQGIKQKSLVELERLANRSSNLNLTNRKST-----DGNNA 467
 QY 464 G--GEYSSENSITVAQOVAYONLRIIRINILEDLSKAWCREQRAALVWNLKSKINPTS 521
 Db 468 THLSNNMESVHNLVYAOQLQFTYDTLRGYINRALAQIAEAWCVDQRTLEVFKLSKINPSA 527
 QY 522 VMSMIYNNRPVSAKRIGDVISVNCIIVDQTSVSLHKSLLSASDEKFCSPRPVTFKFN 581
 Db 528 ILSAIYNKPIARFMDGVGLASCVTINTQISVKVLRDMN--VKESPGRCISRPVVIINFAN 586
 QY 582 DSTIYKQGLGVNNEILLTYYLETQENTYFYFOAKTDMYIYKNYEHKTVPLSSITLTD 641
 Db 587 SSVVQYQGLGEDNEILLGNHRTTECOLPSLKIFIAGNSAYEYVDVLFKRMIDLSSISTVD 646

QY 642 TFIALNFTLENDVFKVIELYTRDEKRLSNVFDIETMREYNVAQVSGRLKDLIDLT 701
 Db 647 SMIALDIDPLENTDFRVLIELSQKRLSSNVFDLEEINRENSYKQRYVEDKVD-- 703
 QY 702 NRNPVDFAFSLMDLDGAVGTVNAVSGVATLFSSIVTGFNFINKPFGGMLMIIVIG 761
 Db 704 PLPPYLKGLDMLSGLAGRAVGAIGAVGAVASVVEGVATFLKNPFGAFTILVAIA 763
 QY 762 VLFALYFLTKTKIYETAPIKWIYPEI-----DKLXE-----R 794
 Db 764 VIIIIYLIYTORRUCMOPLOQLFPYLVASDGTTVTSNGTKDSLQAPPSYEESVNSGR 823
 QY 795 EG-----KSEIAPISEEELERIVLAM-----HIHQONS-----HMETKTR-KDPKDS 835
 Db 824 KGPQPPSDASTAPPYNEAQYQMLALVRLDAEQRAQQNGTSLDQGTGTQDKGQKPN 883
 QY 836 ILTRAQNMRLKRGVSNLKNAESVE 860
 Db 884 LLRLRH---RKNGYRHLKSDDEE 905

RESULT 5
 VGLB_HCMVA STANDARD; PRT; 906 AA.
 AC P06473;
 DT 01-JAN-1988 (Rel. 06, Created)
 DT 01-JAN-1988 (Rel. 06, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Glycoprotein B precursor.
 GN GB OR UL55.
 OS Human cytomegalovirus (strain AD169).
 OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
 OC Betaherpesvirinae; Cytomegalovirus.
 OX NCBI_TaxID=10360;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=87080291; PubMed=3024973;
 RA Cranage M.P., Kouzarides T., Bankier A.T., Satchwell S., Weston K.,
 RA Tomlinson P., Barrell B.G., Hart H., Bell S.E., Minson A.C.,
 RA Smith G.L.;
 RT "Identification of the human cytomegalovirus glycoprotein B gene and
 RT induction of neutralizing antibodies via its expression in
 RT recombinant vaccinia virus.";
 RL EMBO J. 5:3057-3063(1986).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=87151139; PubMed=3029980;
 RA Kouzarides T., Bankier A.T., Satchwell S.C., Weston K.,
 RA Tomlinson P., Barrell B.G.;
 RT "Large-scale rearrangement of homologous regions in the genomes of
 RT HCMV and EBV.";
 RL Virology 157:397-413(1987).
 RN [3]
 RP SEQUENCE OF 735-906 FROM N.A.
 RX MEDLINE=87061230; PubMed=3023690;
 RA Kouzarides T., Bankier A.T., Satchwell S.C., Weston K., Tomlinson P.,
 RA Barrell B.G.;
 RT "Sequence and transcription analysis of the human cytomegalovirus DNA
 RT polymerase gene.";
 RL J. Virol. 61:125-133(1987).
 RN [4]
 RP COMPLETE GENOME.
 RX MEDLINE=90269039; PubMed=2161319;
 RA Chee M.S., Bankier A.T., Beck S., Bohni R., Brown C.M., Cerny R.,
 RA Hornell T., Hutchinson C.A. III, Kouzarides T., Martignetti J.A.,
 RA Preddie E., Satchwell S.C., Tomlinson P., Weston K.M., Barrell B.G.;
 RT "Analysis of the protein-coding content of the sequence of human
 RT cytomegalovirus strain AD169.";
 RL Curr. Top. Microbiol. Immunol. 154:125-169(1990).
 CC -!- SIMILARITY: BELONGS TO THE HERPESVIRUSES GLYCOPROTEIN B FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; X17403; CAA35414.1; -;
 DR EMBL; X04606; CAA28274.1; -;
 DR EMBL; M17209; AAA46009.1; -;
 DR EMBL; M14709; AAA45987.1; -;
 DR PIR; A25365; VGBEC1.
 DR InterPro; IPR000234; Glycoprot B.
 DR Pfam; PF00606; Glycoprotein B; 1.
 DR ProDom; PD000693; Glycoprot B; 1.
 KW Glycoprotein; Transmembrane; Signal.
 FT SIGNAL 1 24 POTENTIAL.
 FT CHAIN 25 906 GLYCOPROTEIN B.
 FT DOMAIN 25 706 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 707 722 POTENTIAL.
 FT TRANSMEM 728 747 POTENTIAL.
 FT TRANSMEM 751 771 POTENTIAL.
 FT DOMAIN 772 906 CYTOPLASMIC (POTENTIAL).
 FT CARBOHYD 37 37 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 68 68 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 73 73 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 85 85 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 208 208 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 281 281 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 286 286 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 302 302 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 341 341 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 383 383 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 405 405 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 409 409 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 417 417 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 447 447 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 452 452 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 464 464 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 465 465 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 554 554 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 585 585 N-LINKED (GLCNAC. .) (POTENTIAL).
 SQ SEQUENCE 906 AA; 102003 MW; D4C7A6A3C7083FEE CRC64;
 Query Match 24.7%; Score 1121.5; DB 1; Length 906;
 Best Local Similarity 30.4%; Pred. No. 2.9e-61;
 Matches 282; Conservative 186; Mismatches 350; Indels 111; Gaps 25;
 QY 14 LWLYQVALYSLSTAETGVTSPNATWSTESPLTGHYGHDSHGGRGNN-ENRD---- 68
 Db 5 IWCL--VVCNLCIVCLGAAVSSSSSTSHATSS-----THNGSHTSRTTSAQTRSVSQ 55
 QY 69 -----SSEQNKNYIG-----SPSTFPYRVCSAGVGDFVRFQTDHVC---- 105
 Db 56 HVTSSSEAVSHRANETIYNTTLKYGDVVGVNTTKYPRVCSMAOQGTDLIRFERNICTSMK 115
 QY 106 PDASDMVHSGILLIYKQNIIPMFVRKRVKVVTTSTVYNGIYSDITNHTYKSIIEP 165
 Db 116 PINEDL--DEGIMVYKRNIVAHFTFKRVYQKVLTFRRSYAYIYTTLLGNSNTYVAPP 173
 QY 166 WETEKMDTIYQCFNSRLNLTGGNLLTYVDRDDI-NMTVFLQPVGVDPVKRVKYSQPELY 224
 Db 174 WEIHHINKFAQCYSSYSRVIGTVFVAYHRDSVENKTMQLIPDDYSNTHSTRYVTVKQW 233
 QY 225 LEFGWFGSYRRRTTNCVCLMDMFARSNPPDFEFVATGDTVEMSPFWSGDEHDKHKE 284
 Db 234 HSRGSTW-LYRETCLNCLMTITARSKYPHYFATSTGDDVVISPYNGNTRNASYFGE 292
 QY 285 KPMFVSVINNYKVY-DYQNRGTVP-LGKTRIFLDREYTLISWE-KHLKNMSYCPDLTWKA 341
 Db 293 NADKFFIFPNYTVISDFGRNPAAPETHRLVAFERADSVISWDIQDEKNVT-CQLTFWEA 351
 QY 342 FYNGIQTEHSGSYHFVANDITASFTSKEDMKENFTTYHCLNEBIKAEIEKKY-AKNST 400

```

Db 352 SERTIREADSHTFSAKTATFLSKQEVNWSDSALDCVRDEAINKLOQIFWTSYNOT 411
Qy 401 HSKYGLDKYPTKDGGLYLVNVOPLQNRLDAKKNLNNETYS---RRSRQAESTDPMME 457
Db 412 YEKYGNVSFETSGGLVVFQGIKQSLVLELRLANRSSLNTHRTTRSTSDNNTHLSS 471
Qy 458 MTGNAGGEVSSSENSITVAQOVAYDNLRIINNILEDKAKVCREQHRALVWNLSKI 517
Db 472 M-----ESVHNLVYAQLQFTYDLRGVYNRALAQIAEAMCVQORRTLEVPKELSKI 522
Qy 518 NPTSVMSMIYNRPVSARKIGDVISVNSCIVVDOTSLSLHSLRLLSADEKCFSRPPTVF 577
Db 523 NPSAIIAISYKPIAFRFGDVLGLASCVTINQTSVKVLDMN-VKESPCRCYSRPVIF 581
Qy 578 KFMNDSTIYKQGLGVNNEILLTTTLETQCENTEYFQAKTDMYIYKNYEHKLTVPULSI 637
Db 582 NFANSSVYVQGLGEDNEILLGNHRTTECOLPSLKIFIAGNSAYEYVDYLFKRMIDLSSI 641
Qy 638 TTLDTFTALNFTLLENVDFKVIELYTRDEKRLSNVFDIETMFREYNYAQRVSLGRKDLL 697
Db 642 STVDSMIALDIDPLENTDFRLEYSQKELRSSNVFDLEIMREFNSYKQRYKVEDKV 701
Qy 698 DLSTNRQVDAFGSLMDDLGAQGVQTVVNAVSGVATLFSISIVTGFINFINKNPFQGLMII 757
Db 702 D---PLPPYLKGLDMLSGLAGAKAVGAIGVAGVAVSVGEVATFLKNPFGAFTIL 758
Qy 758 VVIGLVFAIYFLTKTKIYETAPIKMIYPEI-----DKLKE----- 793
Db 759 VAIATVITLYITFRRLCTOPLQNLFPVLVSADGTTVTSKDTSLQAPPSYEBSVY 818
Qy 794 ---REG-----KSEIAPISEEELERIVLAM-----HIHQONS-----HMETKTP-KD 831
Db 819 NSGRKGCPSPSSDASTAAPPTNEQAYQMLLALRLDAEQRAQOQNGTDSLGGTGTQDKG 878
Qy 832 PKDSILTRAQNLKRSKSYNLKNAESVE 860
Db 879 QKPNLRLRLRH---RKNGYRHLKDSDEEE 904

RESULT 6
VGLB_HSV7J STANDARD; PRT; 822 AA.
AC P52352;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Glycoprotein B precursor.
GN GB OR U39.
OS Human herpesvirus (type 7 / strain JI) (HHV7).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Simplexvirus.
OX NCBI_TaxID=57278;
RN [1]
RP SEQUENCE FROM N.A.
RA Nicholas J.;
RL Submitted (JAN-1996) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: BELONGS TO THE HERPESVIRUSES GLYCOPROTEIN B FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement. See http://www.isb-sib.ch/
CC or send an email to license@sb-sib.ch.
CC -----
CC EMBL; U43400; AAC54701.1; -.
DR PIR; T41941; T41941.
DR InterPro; IPR000234; Glycoprot_B.
DR Pfam; PF00606; Glycoprotein_B; 1.
DR ProDom; PD000693; Glycoprot_B; 1.
DR Glycoprotein; Transmembrane; Signal.
KW

```

```

FT SIGNAL 1 22 POTENTIAL.
FT CHAIN 23 822 GLYCOPROTEIN B.
FT DOMAIN 23 649 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 650 670 POTENTIAL.
FT DOMAIN 671 683 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 684 704 POTENTIAL.
FT DOMAIN 705 822 EXTRACELLULAR (POTENTIAL).
FT CARBOHYD 152 152 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 244 244 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 324 324 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 338 338 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 352 352 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 358 358 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 387 387 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 390 390 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 482 482 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 517 517 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 567 567 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 822 AA; 93147 MW; F2B410BD359C6498 CRC64;

Query Match 23.9%; Score 1084.5; DB 1; Length 822;
Best Local Similarity 33.8%; Pred. No. 4.8e-59;
Matches 272; Conservative 145; Mismatches 316; Indels 77; Gaps 24;

Qy 78 GSPSTPYRVCASGVGDVRFQTDHVCPD-ASDMVHSEGLIILYKQNIIPFMRVRKYR 136
Db 28 GHNOHLPFICSTATGTDLVRDFREVCASGYSNIKTTEGILIIYKTKIEAHTFSVRTFK 87
Qy 137 KVVTTSTVYNGI-----YSDSITNQHTFYKSIETPETERKMDTIYQCFNSURLNTG--GNLL 190
Db 88 KELTFQTYRDVGTVYFLDRTVTI--TLPMPIE--EVHMVNTTEARCLSSIVKSEEEYV 143
Qy 191 TYVDRODIINMTVFLQVDGVTDPVKRYGQPELYLEPGFWGYSRRRTTVNCLEMDMFAR 250
Db 144 AVHKDYVNVKTLDLIPLNFKSDTVRYIITTKPEFLRNGPLW-FYSTSTINCIVTDCIAK 202
Qy 251 SNPPDFVFVATGDTVEMSPFWSGEDDHE-NKMHEKPFVSVVNNY---KVVDYQNRGTV 306
Db 203 TKYPDFFALSTGETVEGSPFYNGINSKTFNEPEKILF---RNNTMLKTFDDGSGKNF 259
Qy 307 PLGKTRIFLDREBYTSLSEKHLKNMSYCPLTLWKAFYNGIQTEHSGSYHFVANDITASFT 366
Db 260 VTLTKMAFLKGNITFSEWQNEESSICLLKHWMTIPHALRAENANSFPIAQLTASPV 319
Qy 367 TSKEDEKMFNTYHCLNEIKABIEKY-AKVNSTHSHKYGDLKYFKTDGGLYLVWQPLIQ 425
Db 320 TGRSNYTLSDSKYNCINSNYTSLDEIYQYVNNSHDKNGSYEIFKTEGDLILTIWQPLIQ 379
Qy 426 NRIILDAKKNLNNETYSRRSRQAEESTTDPWMENTGNGAGGEYSSENSITVAQVQYADNL 485
Db 380 RKTUVLEN-FSNASRRKRR-----ELETNKDIVVQIQLYDYL 418
Qy 486 RIRINNILEDLSKAWCREQHRALVWNLSEKINPTSVMSMIYNRPVSARKIGDVISVNSC 545
Db 419 KDYINTALGKLAEAWCLNQKRTITVLHELKISPSGIIISAVYKGPMSAKLIGDVLAVSKC 478
Qy 546 IVVDQTSVLSLHSLRLLSASDE---KCFSRPPTVFKPMNDS-TIYKQGLGVNNEILLTTT 601
Db 479 IEVNTQSVQLHKSMLRTKDSYDALRCYSRPLTYSFANSKRETYLGQLGLDNEILLGNH 538
Qy 602 YLETC-QENTEYFQAKTDMYIYKNYEHKLTVPULSIITLDTFIALNFTLLENVDFKVI 660
Db 539 RTEECQSNTKIFLSKGF-AHIFKQTYVNSLSLITIEALDAFVDNLIDPLENADFILLE 597
Qy 661 LYTRDEKRLSNVFDIETMFREYNYAQRVSLGRKDLLSLSTNRQVDAFGSLMDDLGA 720
Db 598 LYTKDELKANVDFLETLILREYNSYK---SALHHIETKIATVTPTVIGGIDTFFKGLGAL 654
Qy 721 GQTVNNAVSGV-ATLFSSIVTGFINFINKPFGGLMIIIVVIGLVFAIYFLTKTKIYETA 779
Db 655 G-LGLGAVLGTAGALGDVNVGVFSFLKNPFGGALTLLTGLVIGLVIFLFRHKLAAQT 713
Qy 780 PIKMIYPEIDKLER--EGKSEIAPISEEELERIVLAMHIHQONSHEMTKTRKD--PKDS 835

```


CC -!- SIMILARITY: BELONGS TO THE HERPESVIRUSES GLYCOPROTEIN B FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL; M97928; AAA43847.1; -;
 DR PIR; A44047;
 DR InterPro; IPR000234; Glycoprot B.
 DR Pfam; PF00606; Glycoprotein_B_1.
 DR ProDom; PD000693; Glycoprot_B_1.
 KW Glycoprotein; Transmembrane; Signal.
 FT SIGNAL 1 23 POTENTIAL.
 FT CHAIN 24 830 GLYCOPROTEIN B.
 FT DOMAIN 24 651 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 652 674 POTENTIAL.
 FT DOMAIN 675 688 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 689 709 POTENTIAL.
 FT DOMAIN 710 830 EXTRACELLULAR (POTENTIAL).
 FT CARBOHYD 155 155 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 228 228 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 247 247 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 286 286 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 329 329 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 355 355 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 361 361 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 486 486 N-LINKED (GLCNAC. .) (POTENTIAL).
 SQ SEQUENCE 830 AA; 93266 MW; 44B5B8FC6276DC68 CRC64;
 Query Match 23.1%; Score 1050.5; DB 1; Length 830;
 Best Local Similarity 32.3%; Pred. No. 5.9e-57;
 Matches 244; Conservative 160; Mismatches 293; Indels 59; Gaps 15;
 QY 83 FPYRVCASGVGVFRFQTDHVC--PDASDMVHSEGLLIYKONIIPMFVRKRVKVVTT 141
 DB 36 YPFRICSIAGTDLMLRDRDISCPYKSNKMSGFIILYKNTIETFFPVRYKKELTF 95
 QY 142 STVNGIYSISINQHTFYKSIPEWTEKMDTIYQCNSRLN--TCGNLLTVVDRDIN 199
 DB 96 QSSYRDVGWVFLDRTVMGLAMPVYEANLVNSHAQCYSAVAKRPDCTVFSAFHEDNKN 155
 QY 200 MTPVLPQVDGTPDVKSYGSOPELYLEPCFWGYSRRRTTVNCELMDMPARSPDPFFV 259
 DB 156 NTLNLPFLNFKSINTKRFITTKPEYFARGPLW-LYSTSTSLNCIVTEATAKAKYPSYFA 214
 QY 260 TATGDTVMGSPFWGSD--DHENKMKHPWFSVINNYKV--DYQN--RGTVPLGKTRIFL 315
 DB 215 LTGGEIVGSPFFNGSKHFAEPLEK---LTILENTMIEDLMNGMNGATTLVRKIAPL 271
 QY 316 DREBYTLSEWKLKNSYCYPLTLWKAFYNGIQTEHSGSYHFVANDITASFTTSKEDMKEF 375
 DB 272 EKADTLFSWEIKENESVCMKHWTTVTHGLRAETDTHYFISKELTAAPVAKESLNT 331
 QY 376 NTHYCHLNEELKAEIKY-AKVNSTHYSKYGDLKYFTDGLYLWVQPLONRLLDKWK 434
 DB 332 DPKQTCIKDFEKLINENVYMSDYNDTYSMNGSVQIFKTTGDLILWQPLVQKSLMFL--- 388
 QY 435 LNNETYRRRRROAESTTDPMMETMGAGGEYSSENSITVAOVQYADNLRIRNNILE 494
 DB 389 ---EQGEKIRRRDDV-----DVKSRHDILYVQLQYLYDLYDKYINDALG 431
 QY 495 DLSKAWCREQRAALVNNELSKINPTSMVMYINRPVSAKRIGDVISVNCIVVDQTSVS 554
 DB 432 NLAESWCLDQKRTITMLHELKSKISPSISIVSEVYGRPISAQLHGDVLAISKIEVNQSSVQ 491
 QY 555 LKSLRLLSA-----SDEKCSRPPVTFKFNDS--TIYKGLGVNNELLITTYLETCQEN 609
 DB 492 LKSMRVVDKAGVRSTCMYCNRLPLVTFSEVNSTPEVVPGQLGLDNEILLGDHRTCECEIP 551

QY 610 TEYFQAKTDMYIKYVEHLKTVPLSSITLDTFIALNFTLLENVDFKVIETYRDEKRL 669
 DB 552 SKTIFLSGNHAYVTDYTHNSTPIEDIEVLDFAIRLKIDPLENADFKVLDLSPDELSR 611
 QY 670 SNVFDIETWFRVNYAQRVSGLRKDLDLSTNRNOFVDAFGSLMDDDLGAVGTQVNVAS 729
 DB 612 ANVFDLENILREYNSYK--SALYTIETAKIATNTPSYVNGINSFLOGLGAGTGLGSGVIS 668
 QY 730 GVATLSSIVTGFINKPFGGMLMIVVIGLFAIYFLTKTKIYETAPIKWIYP--- 786
 DB 669 VTAGALGDIVGVGVFLKNPFGGLMLLAIIVVVIIIVVVRQRHVLSKPIDMFFPYAT 728
 QY 787 -----EIDKLEREGKSEIAPISSEE 807
 DB 729 NPVTTVSSVTGTTVWKTPTSKVDGDTGTSVA-VSEKE 763
 RESULT 9
 VGLB HSV6Z STANDARD; PRT; 830 AA.
 ID VGLB HSV6Z AC P36320;
 DT 01-JUN-1994 (Rel. 29, Created)
 DT 01-JUN-1994 (Rel. 29, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 GN Glycoprotein B precursor.
 GN GB OR U39.
 OS Human herpesvirus (type 6 / strain Z29) (HHV6).
 OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
 OC Betaherpesvirinae; Roseolovirus.
 OX NCBI_TaxID=36351;
 RN (1)
 RP SEQUENCE FROM N.A.
 RX MEDLINE=93033157; PubMed=1329336;
 RA Chou S., Marousek G.I.;
 RT "Homology of the envelope glycoprotein B of human herpesvirus-6 and
 RT cytomegalovirus."
 RL Virology 191:523-528(1992).
 CC -!- SIMILARITY: BELONGS TO THE HERPESVIRUSES GLYCOPROTEIN B FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL; M97927; AAA43846.1; -;
 DR PIR; B44047;
 DR InterPro; IPR000234; Glycoprot B.
 DR Pfam; PF00606; Glycoprotein_B_1.
 DR ProDom; PD000693; Glycoprot_B_1.
 KW Glycoprotein; Transmembrane; Signal.
 FT SIGNAL 1 23 POTENTIAL.
 FT CHAIN 24 830 GLYCOPROTEIN B.
 FT DOMAIN 24 651 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 652 674 POTENTIAL.
 FT DOMAIN 675 688 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 689 709 POTENTIAL.
 FT DOMAIN 710 830 EXTRACELLULAR (POTENTIAL).
 FT CARBOHYD 155 155 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 247 247 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 286 286 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 329 329 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 361 361 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 486 486 N-LINKED (GLCNAC. .) (POTENTIAL).
 SQ SEQUENCE 830 AA; 93209 MW; 5E4FE7EC8C43349 CRC64;
 Query Match 22.9%; Score 1040.5; DB 1; Length 830;
 Best Local Similarity 32.3%; Pred. No. 2.4e-56;
 Matches 244; Conservative 155; Mismatches 298; Indels 59; Gaps 15;
 QY 83 FPYRVCASGVGVFRFQTDHVC--PDASDMVHSEGLLIYKONIIPMFVRKRVKVVTT 141

```
Db 36 YPFRICSTAKGDLRFRDIDSCSPYKSNKSEGFFIYKTIETFPVRYTKNELTF 95
Qy 142 STVYNGIYSDSTNQHFFKSETEKMDTIYOCFNSLRN--TGGNLLTYVDRDDIN 199
Db 96 QTSYRDGVVFLDRTVGLMGLVPEANLVSAVIAKRPDGTVFSAHYHEDNNKN 155
Qy 200 MTFVLQPDGVTPDKVKYQSGPELYLEPCGFWGYSYRRRTTNCELMDFARSNPPDFPV 259
Db 156 ETELEFLPFLFKSVTKRFTTKPEYFARGPLW-LYSTSTSLNCIVTEATAKAYFSPFA 214
Qy 260 TATGTVEMSPWASED--DHNKMKHEKWFVSVINNYKV--DYON--RGTVPGLKTRIFL 315
Db 215 LTTGIVEGSPFPDGSNGKHAEPLEK--LTIENYTMIEDLMNGMGATTLVVKIAFL 271
Qy 316 DREYTLSEKHLKMWSCPLTLWKAIFYNGIQTEHSGSVHFVANDITASFTTSKEDMEF 375
Db 272 EKGDTLFSWEIKEENESVCMKXHTVTTHGLRAETDETHYFISKELTAFAVASKESLNT 331
Qy 376 NTTYHCLNEEIKAEIEKRY-AKVNSTHSGYGLKDYKFDGGLYLVQPLIQNRLDANK 434
Db 332 DPQKTCIKNEFEKIITDVMYSDYNDYSNMGYSQIFKTTGDLILLQPLVQKSLM---V 387
Qy 435 LNNETYSRRSROASTTDPMMWMTGNGAGGYSSENSITVAQVQYAYDNLRIRINNILE 494
Db 388 LEQGSWNLRRRRDLV-----DVKSRHDILYVQLQYLYDTLKDYNIDALG 431
Qy 495 DLSKAWCEQHOHRAALVWNLKINTSVMSIMYINRPSVAKRIGDVISVNCIVVDQTSVS 554
Db 432 NLAESWCIDQKRTITMLHLSKISPSISVSEYGRPISAOQLHGDVLAISKIEVNQSSVQ 491
Qy 555 LHKSLRLLSA-----SDEKCFSPPTVPFKPMDS--TIYKGLGVNNEILLTITYLETQEN 609
Db 492 LYKSNRVDAKGVRSCTWYRPLVTFSEVNSTPEVILGQLGDLNEILLGHRTEECIP 551
Qy 610 TERYFOAKTDMYIKYKNEHLKTVPLSSITTLDTFIALNFTLENVDFKVIETRDEKRL 669
Db 552 STKIFLGNHARVYDYTHNSTPIEDIEVLDAFIRLKIDPLENADFLLDLYSPDELSR 611
Qy 670 SNVFDIETMFREYNYAQRVSLGRKDLDLSTNRNQFVDAFGSLMDDLGAVGQTVNVAVS 729
Db 612 ANVFLENILREYNSYK---SALYTEAKIATNTPSYNGINSFPLQGLGAIGTGIGSVIS 668
Qy 730 GVATLFSSIVTGFINFKIPFCGMLMIWIGVLEAIVFLTKTKIYETAPTKMYP---786
Db 669 VTAGALGDI VGVSVFLKNPFGGGLMLAIIVVVIIIVFVRQKVLKSIDMMFPYAT 728
Qy 787 -----EIDKLKEREKSGEAPISEEE 807
Db 729 NPVTTVSSVTGTVVVKTPSVKADGTSVA-VSEKE 763

RESULT 10
VGLB_MCMVVS
ID VGLB_MCMVVS STANDARD; PRT; 928 AA.
AC P27171;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 01-OCT-1993 (Rel. 27, Last annotation update)
DE Glycoprotein B precursor.
GN GB OR UL55.
OS Murine cytomegalovirus (strain Smith).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Betaherpesvirinae; Muromegalovirus.
OX NCBI_TaxID=10367;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92292266; PubMed=1318410;
RA Rapp M., Messerle M., Buehler B., Tanneheimer M., Keil G.M.,
RA Koszinowski U.H.;
RT Identification of the murine cytomegalovirus glycoprotein B gene and
RT its expression by recombinant vaccinia virus.;
RL J. Virol. 66:4399-4406(1992).
```

```
RN [2]
RP SEQUENCE OF 1-280 FROM N.A.
RX MEDLINE=93033129; PubMed=1329325;
RA Messerle M., Keil G.M., Schneider K., Koszinowski U.H.;
RT "Characterization of the murine cytomegalovirus genes encoding the
RT major DNA binding protein and the ICP18.5 homolog.";
RL Virology 191:355-367(1992).
RN [3]
RP SEQUENCE OF 514-928 FROM N.A.
RX MEDLINE=92024072; PubMed=1718083;
RA Elliott R., Clark C., Jaquish D., Spector D.H.;
RT "Transcription analysis and sequence of the putative murine
RT cytomegalovirus DNA polymerase gene.";
RL Virology 185:169-186(1991).
CC -!- FUNCTION: IMPORTANT DETERMINANT OF VIRUS INFECTIVITY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL; M86302; AAA45936.1; -
DR EMBL; X67021; CAA47416.1; -
DR EMBL; M73549; AAA45939.1; -
DR PIR; A41992; VGBEMC.
DR InterPro; IPR000234; Glycoprot_B.
DR Pfam; PF00606; Glycoprotein B; 1.
DR ProDom; PD000693; Glycoprot_B; 1.
KW Glycoprotein; Transmembrane; Signal.
FT SIGNAL 1 36
FT CHAIN 37 928 GLYCOPROTEIN B.
FT DOMAIN 37 739 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 740 755 POTENTIAL.
FT TRANSMEM 761 780 POTENTIAL.
FT TRANSMEM 784 804 POTENTIAL.
FT DOMAIN 805 928 CYTOPLASMIC (POTENTIAL).
FT CARBOHYD 93 93 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 223 223 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 315 315 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 354 354 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 398 398 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 429 429 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 470 470 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 618 618 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 667 667 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CONFLICT 522 522 K -> R (IN REF. 3).
FT CONFLICT 527 527 D -> E (IN REF. 3).
FT CONFLICT 556 556 S -> T (IN REF. 3).
FT CONFLICT 562 562 I -> V (IN REF. 3).
FT CONFLICT 583 583 K -> Q (IN REF. 3).
FT CONFLICT 587 587 S -> K (IN REF. 3).
FT CONFLICT 617 617 I -> V (IN REF. 3).
FT CONFLICT 672 672 E -> D (IN REF. 3).
FT CONFLICT 675 675 D -> Q (IN REF. 3).
FT CONFLICT 806 806 R -> RQRSAAR (IN REF. 3).
FT CONFLICT 890 890 N -> T (IN REF. 3).
SQ SEQUENCE 928 AA; 104109 MW; 0EDBF3FE7A43810D CRC64;

Query Match 22.7%; Score 1034; DB 1; Length 928;
Best Local Similarity 31.3%; Pred. No. 7.le-56;
Matches 264; Conservative 155; Mismatches 324; Indels 100; Gaps 25;

Qy 8 RGSVLALWLYQVAL-----YLSIAETGVTSF-----PNTATWSTESPLT 48
Db 7 RGRSSWYAMSTALAVTIWCLIACTSEVIAASTPTGTTTPAKTDTTSETSAETATS 66
Qy 49 GHYTHDSSHGGRNNENRDS-----EQKNKIYSGSPFPYRVCSASGVGVDF 97
Db 67 G-----AATGKKEATPTQASKITGTTIVFVNEDMVSVDIDKYPRVCMAVST-DLV 119
```

```
QY 98 RFQDTHVCPDASDMVH-SEGILLIYKQNIIPMFVRVKYKVVTTSTVYNGIYSDSITNQ 156
|||
Db 120 RFKSIDCINHPTKPTQVQEGIMVYKENIVAYFEVITYHK-----DAIFORSADITINY 175
|||
QY 157 H---TFYKSHIP-WETEMKPTIYQCFNSL-RLNTGGNLLTYVDRDDINMTVFLOPQDGV 211
|||
Db 176 FLGTSVTKMAFFIWELDEVNRNRCYSAASRIINGEVVYVAYHEDSYRNYTWLVEDDYRS 235
|||
QY 212 PDVKRYGSQLYELPQWFGWSYRRRTTUNCELMDMFARNPDPFPVATGDTVMESP 271
|||
Db 236 KNSKRYVTTSRHYKGAWTW-RYTESCNMNCVVVTKARGNTPYEFFVLSSGEVVEISPF 294
|||
QY 272 WSGEDDHENKMKPEWFWFVSVNNYKVDY-----QNRGTVPGLKTRIFLDRREEYTLWSWK 326
|||
Db 295 YNGENSFPFEDTRNFWR--KNTMTKYFGELAAPKRVLP---MAFLEREDMTIGWEI 349
|||
QY 327 HLKMSYCPCLTWKAFYNGIQTEHSGSYHYFVANDITASFTTSKEDMKFEFT-----T 378
|||
Db 350 FPKQNVTCDWKQWTSRAIRTDNTCYHFVSKGLTATFVASKRKI-DYNTTTEGKNYNT 408
|||
QY 379 YHCLNEBEIKABIEKYYA-KVNSTHYSKYDGLKYFTDGGGLYLVWQPLIQNRLLDKAKN 437
|||
Db 409 FRCVYDFEVEVNRVFEDEYNETHVKDGELEWVTTGGLIVLMQG-----LKA KSLHNL 462
|||
QY 438 ETSYRRRRROAESTTDPW-MEMTNG-----AGGEYSSEN-----SITVAOVQVAYDNL 486
|||
Db 463 EKFAALNNVSATASPPTTAATENGTVTSRRKRSPDNLDVVVDTSYAOLOFTYDVLK 522
|||
QY 487 IRINILEDLSKAWCREQHRRAALVWNLKSNKINPTSMVMIYNRPVSRAKRGDVISVNCI 546
|||
Db 523 DYINDALRNIMDAWCRQKRTAEMLKELSNPSNLSAIYERPVTKLAGDIANSECV 582
|||
QY 547 VVDTSVLSKLSRLLSASDEK---CFSRPVPVTKFNDSTIYKQGLGVNNEILLTTTLY 603
|||
Db 583 KVDQSSVKVLKDMRIF--QDKVVCYNSRPLVWPFQFINSTKLESQGLGENNEIMLGFRT 640
|||
QY 604 ETCQENTYFQAQTMVYKNYEHLKTVPLSSITLDTFLALNFTLENVDFKVIELYT 663
|||
Db 641 ENCDTNRKIFVQVGVYGYRDRFRNVTSLEHIDLVDTLGLDIEPLENTDFKVLLEY 700
|||
QY 664 RDEKRLSNVDFIETMFREYNYYAQRVSGRLKRLDLSTNRNQFVDYAFGLSMDLGVAGOT 723
|||
Db 701 KGLRASNVFSLDEIMREYNSQKHIRTSAKND---NTPSYLLGLDFTMQGLGVAGK 757
|||
QY 724 VNAVSGVATLFSISIVTGFINFKPFGGMLIIVIGVLPAIYFLTKTKIYETAPIKM 783
|||
Db 758 IGVAIGAAGVGVAVSVNAVVTGFLTNPFPGFTTILLVIGLAVVLIIF-----TRPEY 810
|||
QY 784 IYP 786
|||
Db 811 FFP 813
|||

RESULT 11
VGLB_HSVT2
ID_VGLB_HSVT2 STANDARD; PRG; 944 AA.
AC QWRUL5;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Glycoprotein B precursor.
GN GB.
OS Herpesvirus tupaia (Strain 2) (THV-2).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Betaherpesvirinae.
OX NCBI_TaxID=132678;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=9319892; PubMed=10392721;
RA Bahr U., Springfield C., Tidona C.A., Darai G.;
RT "Structural organization of a conserved gene cluster of Tupaia
herpesvirus encoding the DNA polymerase, glycoprotein B, a probable
```

```
RT processing and transport protein, and the major DNA binding protein.";
RL Virus Res. 60:123-136(1999).
CC - SIMILARITY: BELONGS TO THE HERPESVIRUSES GLYCOPROTEIN B FAMILY.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement. See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch.
CC -----
DR EMBL; AF084543; AAD42935.1; --
DR InterPro; IPR000234; Glycoprot B.
DR Pfam; PF00606; Glycoprotein_B; 1.
DR ProDom; PD000693; Glycoprot_B; 1.
KW Glycoprotein; Transmembrane; Signal.
FT SIGNAL 1 ? POTENTIAL.
FT CHAIN ? 944 GLYCOPROTEIN B.
FT DOMAIN 2 761 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 762 782 POTENTIAL.
FT TRANSMEM 783 785 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 807 944 POTENTIAL.
FT CARBOHYD 114 114 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 238 238 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 369 369 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 409 409 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 414 414 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 426 426 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 481 481 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 485 485 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 620 620 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 944 AA; 106059 MW; 1B063A4CD1561EE2 CRC64;

Query Match * 22.2%; Score 1008; DB 1; Length 944;
Best Local Similarity 31.4%; Pred. No. 2.9e-54;
Matches 260; Conservative 161; Mismatches 322; Indels 84; Gaps 25;

QY 15 WYLYQVA-----LYSLIAETGVSPPTATWSTESPLTGHYGHDSHGGRGNENR 67
|||
Db 41 WYATLIASLVWYPTVSSITLTATVSSDTGGAGQASG--GGGGAGDSTPSE-SPETSA 97
|||
QY 68 DSEEQNKNYG-----SPSTFPYRVCSASGVDFERFQTDHVC-PDASDMVHSGILL 119
|||
Db 98 DITVPRVGTGTEWNSNLTSERYPYRICMSOGTDIVRFARTITCAPYDAKSVSTGIML 157
|||
QY 120 IYQNIIPFMFRVKYKVVTTSTVYNGIYSDSITNQHTFYKSIETPEWETEKMDTIYQCFN 179
|||
Db 158 IYKANIVPTFDVFTYQKELFFORSYAYIYTTVLLGNSREHVALPLWEVDAAINWVCYS 217
|||
QY 180 SURLNTGGNLLTYVDRDDI-NMTVFLOPQDGVTPDVVKRYGSQLYELPQWFGWSYRRRT 238
|||
Db 218 SYVRVTIGTEQYVSYHQDSYRNETHMLIPEEYOSGNTRRYVTVKDQVHYVGSTW-LYKETC 276
|||
QY 239 TVNCELMDMFARNPDPFPVATGDTVMESPW---SGEDDHEN--KMKHEKFWFVSVIN 293
|||
Db 277 SMCIVTQTKAKSKYDYDFALSSGLVVEASPPYDTVNGHTTHENRRKHF-----VRE 329
|||
QY 294 NYKVVDYQNRGTVPGLKTRI-----FLDREEYTLWE-KHLKNMSYCPCLTWKAFYNGIQ 347
|||
Db 330 QYRMLERFCAVNAVPV---RVVPKMAFLERPDIVLAWELKEPKNVT-CHLALWETVNAIR 395
|||
QY 348 TEHSGSYHYFVANDITASFTTSKEDMKFEFTT-YHCLNEBEIKABIEKYYAKVNSTHYSKY-- 404
|||
Db 386 TEHATSFHFVSRLTATFVTAKANETLYNNSRYDCIRDSANRTIDRVFRE--EYDQYEL 443
|||
QY 405 -GDLKYFKTDGGLYLVWQPLIQNRLLDKAKGL-----NNETYSRRSRROAESTTDPWME 457
|||
Db 444 DGDVPIFTNGGUTVWVQGLRQ-KALAAALAGIPGANGTTNHSRHRD----- 492
|||
QY 458 MTGNGAGGEYSSENSITVAQVQVAYDNLRIINILEDLSKAWCREQHRRAALVWNLSKI 517
|||
```

Db 493 -TAAIAREHASD--LTYAQLOFAYDTIRDVYNAQIGHIAEAWCUEQORGTGEMLHLSKI 549

QY 518 NPTSMVMYINRPSVSAKRICDVISNCICVWDQTSVSLHSLRLLSAS-----DEKCF 570

Db 550 NPTSMYALYDRPIARLAGDIALAKCEVVDQTVQVORDMRKPKETSVDGTTEQOQFCY 609

QY 571 SRPPVTFKPMNDSTIYKQGVNNEILLTTVLETCQENTYFFQAKTDMYIKYNEHLK 630

Db 610 SRPVVLFREVNSETQYQGLGDNELTGTFRTEACQLPSLKIFVAGKVAYEVRDYLKR 669

QY 631 TVPLSITLDTFIALNFTLLNENVPKVELYTRDEKRLSNVDFIETMREYNYAQRVS 690

Db 670 QIDLOSIDVNTMISLUKPLENTDFQVLELYSRGELKSANVFLEDDIMREYNAKRLR 729

QY 691 GLRKDLDLSTNRNFVDFAFGSLMDLGAQVQ---TVNNAVSGVATLFSSIVTGFNFIK 747

Db 730 YITSKIYNPIP---PFMRGLDDFMSGLGAAGKGLGLVLGAVGAV---ASVVGGSFPT 783

QY 748 NPFQGLMLIIVVIGLVFAIYFTKTKIYETAPIKMIYPEIDKLER 794

Db 784 NPFGLSLIILVAVVIVFLLYQORSQSAVRQPLDFFFPYLAQQTOR 830

RESULT 12

VGLB_GPCM

ID VGLB_GPCM STANDARD; PRT; 901 AA.

AC Q69024;

DT 15-DEC-1998 (Rel. 37, Created)

DT 15-DEC-1998 (Rel. 37, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Glycoprotein B precursor.

GB GB OR UL55.

OS Guinea pig cytomegalovirus (strain 22122 / ATCC VR682) (GPCMV).

OC Viruses; daDNA viruses, no RNA stage; Herpesviridae;

OC Betaherpesvirinae; Muromegalovirus.

OX NCBI_TaxID=103920;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=94279138; PubMed=8009831;

RA Schleiss M.R.;

RT "Cloning and characterization of the guinea pig cytomegalovirus

RT glycoprotein B gene.";

RL Virology 202:173-185(1994).

CC -!- SIMILARITY: BELONGS TO THE HERPESVIRUSES GLYCOPROTEIN B FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch).

CC -----

DR EMBL; L25706; AAA43831.1; -

DR InterPro; IPR000234; Glycoprot B.

DR Pfam; PF00606; Glycoprotein B; 1.

DR ProDom; PD000693; Glycoprot B; 1.

KW Glycoprotein; Transmembrane; Signal.

FT SIGNAL 1 34 POTENTIAL.

FT CHAIN 35 901 GLYCOPROTEIN B.

FT DOMAIN 35 713 EXTRACELLULAR (POTENTIAL).

FT TRANSMEM 714 734 POTENTIAL.

FT TRANSMEM 737 757 POTENTIAL.

FT DOMAIN 758 901 CYTOSOLASMIC (POTENTIAL).

FT CARBOHYD 53 53

FT CARBOHYD 60 60 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 66 66 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 197 197 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 270 270 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 289 289 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 328 328 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 372 372 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 398 398 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 406 406 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 436 436 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 537 537 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 571 571 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 623 623 N-LINKED (GLCNAC. . .) (POTENTIAL).

SQ SEQUENCE 901 AA; 102239 MW; 43ABCA7D45A197AA CRC64;

Query Match 22.0%; Score 999; DB 1; Length 901;

Best Local Similarity 29.1%; Pred. No. 9.6e-54;

Matches 271; Conservative 159; Mismatches 356; Indels 144; Gaps 25;

QY 24 SLSIAETGVTSPNTATWSTESPLTGHYTHDSSHGERGNENRDSSEONKNI-----YG 78

Db 24 ALTILYLGVCSTTPTTPTTVEEDTTVSGNH--SDASRNNVTIR-----NLTSVDF 74

QY 79 SPSTPYVCSAGVDVFRQTDHVCDDASDMVHS-EGILLIYKQNIIPFMRVRYKPK 137

Db 75 QKLYPYRICSMGTDLVRFARTIQCVFPRVNSEEGIMLIYKRNILPYVFTAYTQK 134

QY 138 VVTSTVYNGIYSDITNOHTFYKSIPEWETEKMDTIYQCFNSLRNLNTGGN-LLTVVD 196

Db 135 ELLFORSYKYTYDYLLGYSREFVALPMWEFLVNSRGQCYTSHORVIGADRYIAHND 194

QY 197 DINMTVFLQPDVGTVDKRYGSOPELYLEPGWFGSYRRRTTNCMLDMFARNPPD 256

Db 195 EVNETMMLMRDDMGNDTTRYITVKEHARTPGSVW-LYKETCSMNCIVTKTKGSKFPYD 253

QY 257 FVATATGDTVMSPFWSGED-----DHENKMKHEKPFVSVINNYKVW-DYQNRGTVP 310

Db 254 MFVLPSGVIVNISPFYNSNGKTPREQREKPH-----IWSNYSILKDFGRAL- 302

QY 311 TRI-----FLDREYVLSWEKHKMSYCPILTWKAFYNGIQTEHSGVHFVANDITASF 365

Db 303 ARIYPKMAFYEREDVIGWEVNDGNTVCEMILMETVDRAIRTEYENAFHYVARTLTSTF 362

QY 366 TTSK-----EDMKEFNTYHCLNEEIKAEIEKKYAK-VNSTHSGYKGLKYFKTDGG 415

Db 363 VENKYSPPNNLTEDDIK-----CFKNDACKIEEVLFRDYNEDYMDGNATYHVTGG 415

QY 416 LYLWQPLIQRLDQAKNLNETHYR--RSRQAEESTTDDMMETGNGAGGEYSSESI 473

Db 416 LVIVWQGLKQKSLKALEIAANESAVSATGNSRRKSLPD---ESTGD-----I 461

QY 474 TVAQVOYAYDNLRIINILEDLSKAWCREQHRALVWNLKINPTSVMSIYNRPVSA 533

Db 462 SYAQQFAYDLYRYINQALGHIAEAWCLOKQRTAEVLHLSKINPILLSAIFGVVAA 521

QY 534 KRIGDIVSNCSIVVDQTSVSLHSLRLLSASD--EKCFSRPPVPTFKPMNDSTIYKQGLG 591

Db 532 RVGVDVISLAKIEVNQSTVLIKGMKRFSDGKLEGCSRPVWVWFSMKNSTEVRLQGLG 581

QY 592 VVNEILLTTTLETQENTYFFQAKTDMYIKYNEHLKTVPLSSITLDTFIALNFTLL 651

Db 582 EDNEILLGTHRMETCQTDYRIFVAGDYGVEFQQVFTKINLSEIDIIDTMIALKTEPL 641

QY 652 ENVDPKVIELYTRDEKRLSNVDFIETMREYNYAQRVSGRLKOLLDLSTNR--NQVDA 709

Db 642 ENIDFKVLELYSRDELAQANVFDESIMREYNYOKKR-----LDFVVERVINPIPPA 693

QY 710 FGSL---MDDLGAVGQTVVNAVSVATLFFSSIVTGFNFIKPNFGGLMLIIVIGVLF 766

Db 694 LKGLDEMNGMGAIGKIGEAVGAVGAIGSFGALVTFVTNPFAGFVVFVFCVGCITLV 753

QY 767 YFLTKTKIYETAPIKMIYPEIDK-----LKEREKSEIAP----- 802

Db 754 ITVVRORRAMQRPDFFFPYASQTITSSVADSSIAVAYPGEGTSGDAPPYGEADYG 813

QY 803 -----ISEELERIVLAWHIHQONSHMETKTRDKPKDSILTR- 839

Db 814 YKDLVSADTRVSSSSAGAGADFNEDAVRMLRA--IKRLDDKKRQETEKSSKDSASKN 871

QY 840 -----AQNMLKRCGYSNLKNAESVEM 861

```

Db      872 SETRRPGIMDLRRRRGGYQKLNTEDDVHV 901
RESULT 13
VGLB_HSVMD STANDARD; PRT; 920 AA.
AC Q0464;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DE 16-OCT-2001 (Rel. 40, Last annotation update)
DE Glycoprotein B precursor.
GN U127.
OS Herpesvirus saimiri (type 1 / strain MV-5-4-PSL) (Marmoset
OC herpesvirus)
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae.
OX NCBI_TaxID=10353;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93228440; PubMed=8385913;
RA Eberle R., Black D.;
RT "Sequence analysis of herpes simplex virus gB gene homologs of two
RT platyrrhine monkey alpha-herpesviruses.";
RL Arch. Virol. 129:167-182(1993).
CC -1- SUBUNIT: DIMER, PROBABLY LINKED BY DISULFIDE BONDS.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- MISCELLANEOUS: GB IS THE ONLY GLYCOPROTEIN THAT IS KNOWN TO BE
CC REQUIRED FOR VIRAL GROWTH.
CC -1- SIMILARITY: BELONGS TO THE HERPESVIRUSES GLYCOPROTEIN B FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; M95786; AAA43841.1; -
DR InterPro; IPR000234; Glycoprot B.
DR Pfam; PF00606; Glycoprotein B; 1.
DR ProDom; PD000693; Glycoprot B; 1.
KW Glycoprotein; Transmembrane; Signal.
FT SIGNAL 1 28
FT CHAIN 29 920 GLYCOPROTEIN B.
FT DOMAIN 29 739 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 740 755 POTENTIAL.
FT TRANSMEM 761 781 POTENTIAL.
FT TRANSMEM 784 804 POTENTIAL.
FT DOMAIN 805 920 CYTOPLASMIC (POTENTIAL).
FT CARBOHYD 98 98 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 119 119 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 152 152 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 409 409 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 441 441 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 583 583 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 733 733 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 920 AA; 101661 MW; 62C4D0315D8C2DB7 CRC64;
Query Match 19.8%; Score 900; DB 1; Length 920;
Best Local Similarity 29.0%; Pred. No. 1.2e-47;
Matches 243; Conservative 154; Mismatches 359; Indels 82; Gaps 21;
Qy 60 ERGNENRDS-----EEQNKVIYGSPTFPYRVCSAGVGDFRQTDHVCPD-ASDVMHS 114
Db 95 QRANKTRDARAQLRVSQRIRAEANATSMFVCPPTPGATVQVQFEPRPCPDVAAGKNT 154
Qy 115 EGILLIYKQNLIPMFVRVRYKVVTTSTVY-----NGIYSDSTNTHQTFKSLIEPW 166
Db 155 EGIAVIFKRNIAPKYFATWYKEITVTQWQSGRYLQLTGLYND-----RAPVPF 205
Qy 167 E--TEKMDTITQCFNSLRNLNTGNNLLTYVDRDDINMTVFLQPVGDVTPDVKRYGSOPELY 224

```

```

Db      206 EBITDVINAKGLCRSDVTYVRSORRVAYDRDEWGREVKLVPSKSTSPNSRGWYTTDRMY 265
Qy      225 LEPGWFGSYRRRTTUNCCLMDMFARSNPPDFEFVTATGDTGVMSFWS-GEDDH-ENKM 282
Db      266 -APNAHAGFYKAGTTNCVIEVEARSAYPSNFVLATGDFVYVSPFYGLGEDAHRREYNA 324
Qy      283 HEKPFVSVINNYKVVDYQNRGTVPLGKTRIFLDREBEYTLSEKHLKNMSYCPPLTLWKAF 342
Db      325 YSADRFPQV-DGFFPRDLDSGETAPBPVVRNLTTPKFTIGDWKPKDPSCVSKVQEV 383
Qy      343 YNGIOTEGSHGSYHFVANDITASFTTSKEDMKFNTTYH-----CLNEEIKAEIEKK 393
Db      384 EEMWRAEYGSTPRFTSSLSATFTT-----NVTOYPPORTIELSDCVAREAAVAIDAI 435
Qy      394 YA-KVNSTHSKYGLKYFKTDGGLYLVWQPLIONRLLDKKNLNNETYSR-----RSRQA 448
Db      436 YARRYNASHVKVGLQYLAQGGFLVYQPLISLSLA-----EMYLEAEARALEPA 487
Qy      449 ESTTDPMMEMTNGAGGEYSSENSITVAOVAYDNLIRINNILEDLKAWCREQHRAA 508
Db      488 PLPTTAPAEA--GSRGTLSTTOSVEFARLQFTYDHIQKHNEMLGRIAAAMCOLNQEL 545
Qy      509 LVWNELSKINPTSVMSMIYNRPVSAKRIGDVISVSNICVVDDQTSVSLHLSRLLSASDEK 568
Db      546 VLNWEARKLPNAIASATVGRVGRMLGDVMAVSTCIAVAPHNVIMQNSMR-LPARPKT 604
Qy      569 CSRPVPTFKFMDSTIYKGLGVNNEILLTITTYLETQENTYFYQAKTDMYIY-KNYE 627
Db      605 CYARPLVSFRYADEGELIEGOLGEDNEIRLEQNLEPCTVGHKRYF-VFGDGVYFEEYA 663
Qy      628 HLKTVPLSSITLDTFIALNFTLLENVDFKVIETRYDEKLSNVFDIETMPREYNYAQ 687
Db      664 YSHQVSRADVVPVSTFVDNLNLMLEDEHFLPLEVYTRQEIKDSCLLDYAEVQRNQMHAL 723
Qy      688 RVSGLRKDLDLSTNRNQFVDAFGSLMDDLGAVGQTVVNAVSGVATLFSIVTGFINFIK 747
Db      724 RFSDI--DHIINDTTAALMDGLFRFFDGLGAAGQAIKAVLGAVTEAVISWVGVSFLS 781
Qy      748 NPGGMLMIIVVIGLVFAIYFLTKTKIYETAPIKMIY-----EIDKLKERE 795
Db      782 NPEGALAVGLVLGLTAFAFFALRYIMLRANPMRALYPTTHGIKAEAKASLASGEPRS 841
Qy      796 GKSEIAPISEELERIVLAWHIHQNSHMETKTRKPKDSILTRAQN-----MLRKS 848
Db      842 GPGGIEDFEAKLEAEARTMIKYMTLVSAMERTAKAKRGTSARISRLHLDVMVRKN 899
RESULT 14
VGLB_HSVMD STANDARD; PRT; 865 AA.
AC F18538;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Glycoprotein B precursor.
GN GB.
OS Marek's disease herpesvirus (strain RB-1B) (MDHV).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Marek's disease-like viruses.
OX NCBI_TaxID=33707;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89293086; PubMed=2544666;
RA Ross L.J.N., Sanderson M., Scott S.D., Binns M.M., Doel T., Milne B.;
RT "Nucleotide sequence and characterization of the Marek's disease
RT virus homologue of glycoprotein B of herpes simplex virus.";
RL J. Gen. Virol. 70:1789-1804(1989).
CC -1- SIMILARITY: BELONGS TO THE HERPESVIRUSES GLYCOPROTEIN B FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way

```

CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL; D13713; BAA02866.1; -
 DR InterPro; IPR000234; Glycoprot_B.
 DR Pfam; PF00606; Glycoprotein B; 1.
 DR ProDom; PD000693; Glycoprot B; 1.
 KW Signal; Glycoprotein; Transmembrane.
 FT SIGNAL 1 21 POTENTIAL.
 FT CHAIN 22 865 GLYCOPROTEIN B.
 FT DOMAIN 22 865 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 683 700 POTENTIAL.
 FT TRANSMEM 709 729 POTENTIAL.
 FT TRANSMEM 732 752 POTENTIAL.
 FT DOMAIN 753 865 CYTOPLASMIC (POTENTIAL).
 FT CARBOHYD 184 184 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 332 332 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 364 364 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 406 406 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 425 425 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 631 631 N-LINKED (GLCNAC. .) (POTENTIAL).
 SQ SEQUENCE 865 AA; 98091 MW; B30E93CIAC65C6C3 CRC64;
 Query Match 19.3%; Score 876.5; DB 1; Length 865;
 Best Local Similarity 27.0%; Pred. No. 3e-46;
 Matches 231; Conservative 164; Mismatches 386; Indels 75; Gaps 16;
 QY 51 YGTHDSSHGCRGNENRSEONKNIYGSPTFPYRVCSASGVGVFRFQTHVCPDASD 110
 DB 17 YGTNSP-----STQNVTSREVVSVQVLSSEESTFYLCPPVPGVSTVIRLEPPKCPERK 71
 QY 111 MVH-SEGILLYKONIIPFMFRVRYKRVVTTSTVYNGIYSDSIINQHTFYKSTIEPWS-T 168
 DB 72 ATWEGEITALFENISYKFKVLYKNIQTWTGTTRQINRTDTRPVSIEIT 131
 QY 169 EKMDIYQCFNSRLNTGNLLTYVDRODINMTVELQPDVGPDKRYGSOPELYLEPG 228
 DB 132 DLIDKGRCSKARYLRNVVVEAFDRDAGEKQVLLKPSKFTPSRAWHTTNETYT--- 188
 QY 229 WFMGS---YRRRTVNCMLDMFARNPPFPFFVTATGDTVEMSPFW--SEGDHENQKH 283
 DB 189 -WVGSPWIRYGTSTVNCVIEEDARSVPFYSYFANANGDIANISPFGLSPPEAAEPMG 247
 QY 284 EKPWFVSVINNVYDQNGRVPLGKTRIFLDREYTLSEKHLKNNISCPPLTWKAFY 343
 DB 248 YPDNFKQLDSYFSDLDKRRKASLPVKNFLITSHFVGVGDWAPKTRVCSMTWKKEVT 307
 QY 344 NGIOEHSYHFVANDITAGFTTSKEDMKENFTY-----HCLNEEIKAEIEKY 394
 DB 308 EMLRATVNGRYRFMARELSATFIS-----NTTEPDNRIILQCICKRAEAAIEQIF 359
 QY 395 -AKVNSTHSGYGLKFKYTDGGLYLVWQPLQ-----NRLLDANKNLN 436
 DB 360 RTKYNDSHVKGVHGVYFLALGGFVAYQPVLSKSLAHMYLRMLRDNRTDMLDLVNNKH 419
 QY 437 -----NETYSRRSRQAESTTDPMMETGNGAGGYSSENSITVAQVQYADNLRIRNN 491
 DB 420 AIYKKNATSLSLRDIRNAPNRKITLDDTTA---IKSTSSVQFAMQLFDHDIQTHIND 476
 QY 492 ILEDLSKAWCREQHRALVWNLKINPTSVMSMIYNRPVSAKRIGDVISVNCIVVDQT 551
 DB 477 MFSRIATAWCELQNLRELWHEGINKINFSATASATLGRRAAKMLGDAVAVSSCTAIDAE 536
 QY 552 SVSLHKSRLLSASDEKFSRPVTFKFMNDSTIYKGLGVNNELLTTTTFLETCQENTE 611
 DB 537 SVTLQNSMRVIT-STNTCYSRPLVLFSGENQGNITQGLGENNELLPLEAVEPCSAHR 595
 QY 612 YFQAKTDWYIKVNYEHLKTVPLSSITTLDTFIALNETLLENVDKVELYTRDKRLSN 671
 DB 596 RYFLFGSGYALFENFNFVMDAADIQIASTTFVELNLTLLDBREILFUSVYTKBELRDVG 655
 QY 672 VDIETMFREYNYAQRVSGLRKDLLDLSTNRNQFVDAFGSLMDDLGAVGQTVVNAVSGV 731

DB 656 VLDVAEVARRNQLHKLKFDYDINK-VIEYDTN-YAFMNLGAELEFNGMGQVQAIGKVVYGA 713
 QY 732 ATLFSSIVTGFINFKPFGGMLMIIVVIGVLFALYFLTKTKIYETAPIKMIYPEIDKL 791
 DB 714 AGAIVSTISGVSAFMSNPFAGALGIIIIAGLVAFLAYRYVVKLKNPMKALYPMPTTEV 773
 QY 792 KERECKSEIAPISBEELER-----IVLAMHIHQONSHMETKTRKDPKDSILT 838
 DB 774 LKAQATRELHGEESDDLERTSIDERKLEAEAREMIKYMALVSAEBHEKLRKRGTTAV 833
 QY 839 RAQNMLRKRSYSNLK 854
 DB 834 LSDHLAKMRKIKNSPK 849

RESULT 15

VGLB_HSVBC STANDARD; PRT; 932 AA.

ID VGLB_HSVBC
 AC P12640;
 DT 01-OCT-1989 (Rel. 12, Created)
 DT 01-OCT-1989 (Rel. 12, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Glycoprotein 1 precursor (Glycoprotein GVP-6) (Glycoprotein 11A)
 DE (Glycoprotein 16) (Glycoprotein G130) (Glycoprotein B).
 GN GI OR UL27.
 OS Bovine herpesvirus type 1 (strain Cooper).
 OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
 OC Alphaherpesvirinae; Varicellovirus.
 OX NCBI_TaxID=10323;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=88300884; PubMed=2841484;
 RA Whitbeck J.C.; Bello L.J.; Lawrence W.C.;
 RT "Comparison of the bovine herpesvirus 1 gI gene and the herpes
 simplex virus type 1 gB gene."
 RL J. Virol. 62:3319-3327(1988).
 CC -!- SUBUNIT: DIMER, PROBABLY LINKED BY DISULFIDE BONDS.
 CC -!- SIMILARITY: BELONGS TO THE HERPESVIRUSES GLYCOPROTEIN B FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).

EMBL; M21474; AAA46055.1; -
 EMBL; Z78205; CAB01598.1; -
 EMBL; AJ004801; CAA06106.1; -
 PIR; A28877; VGBEBC.
 DR InterPro; IPR000234; Glycoprot B.
 DR Pfam; PF00606; Glycoprotein B; 1.
 DR ProDom; PD000693; Glycoprot B; 1.
 KW Glycoprotein; Transmembrane; Signal.
 FT SIGNAL 1 67
 FT CHAIN 68 932 GLYCOPROTEIN I.
 FT DOMAIN 68 758 EXTRACELLULAR.
 FT TRANSMEM 759 827 POTENTIAL.
 FT DOMAIN 828 932 CYTOPLASMIC.
 FT CARBOHYD 105 105 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 153 153 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 441 441 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 483 483 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 640 640 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 706 706 N-LINKED (GLCNAC. .) (POTENTIAL).
 SQ SEQUENCE 932 AA; 101195 MW; 9DCEAA85C5FC3DA3 CRC64;

Query Match 18.8%; Score 857; DB 1; Length 932;
 Best Local Similarity 27.5%; Pred. No. 5.2e-45;
 Matches 250; Conservative 161; Mismatches 425; Indels 72; Gaps 21;

```
QY 2 AGSLKRGSLALWLYQVALYSLSIAETGVT-----SPPNTATWSTESPLTGHY 51
Db 111 MVHSEGLLLYKONIIFPMFRVKYRKVVTITSVYNGIYSDSITNQHTFFKYSIEPWE-TE 169
QY 41 AGGARAALAAALLWATWALLAPACRPATTAPPPEEAAAPAPPASPSPGPGDDDA 100
Db 52 GTUDDSHGEGNNENRDEBQNKNIYGSPPFPYRVCASGVDGVRFFOTDHHVCPD-ASD 110
QY 101 ASPDNSTDVRAALRLAQAAGENSRRF-----VCPPPSGATVRLADARPCEYGLG 151
Db 111 MVHSEGLLLYKONIIFPMFRVKYRKVVTITSVYNGIYSDSITNQHTFFKYSIEPWE-TE 169
QY 152 RNYTEGIVGIVKENIAPYTFKVIYKNIIVITWAGSTYA-AITNQYTDVRVPVGMGEITD 210
Db 170 KMDTIYOCFNSLRNLNTCGNLLITVDRDDINNTVFLQPDVGTVDVKRYGSOPELYLEPGW 229
QY 211 LVDKWRCLSKAEVLSGRKVVAFDRDDDPWEAPLKPARLSAPGVRGWHTTDDVYALG- 269
Db 230 FWSGRRRTTNCGLMDMFARSNPPDFFTVATGDTVMESPFW-SGDDHENKMKHPWF 288
QY 270 SAGLYRTGTSVNCIVEEARSVPYDSFALSTGDIITMSPFYGLREGAHRHTSYSPER 329
Db 289 VSVINNYKVVDYQNRGTVPLGKTRIFLDREYTLNWEKHLKNMSYCPDLTLWKAFYNGIQ 348
QY 330 FOLEGYKXKDMATGRKLEKPEVSRNFRTOHVTAMDVPKRNKVCSLAKWREADEMLRD 389
Db 349 EHSGSYHFVANDITASFTTSKEDMKFNNTY-HCLNEEIKAEIEKKY-AKVNSTHSGYGD 406
QY 390 ESRGNFRFTARSLSATFVSDSHTFALQNVPLSDCVIEEAAVERVYREYNGTHVLSGS 449
Db 407 LKYFKTDGGYLVWQPLIONRLDANKL-----NNETYS-----RRSRROAE 449
QY 450 LETYLRGGFWAFRPMLSNEL-AKLYQLELARSNGTLBGLFAAAAPKPGPRARRAAP 507
Db 450 STTDPNMWMTGNG-----AGGEYSSENSITVAQOVAYDNLRINNIILEDLSKAWRE 503
QY 508 SA--PGSGAANGPAGDGDAGGRVTTVSSAEFAALQFYDHIQDHVNTMFSRLATSWCLL 565
Db 504 QHRAALVWNELSKINPTSVMSMIYNRVPSAKRIGDVISVNCIVDQTSVSLHKSRLLS 563
QY 566 QNKERALWAEAKLNPSAASALDRRAAARMGLGDAMAVTYCHELGEGRVFIENSVR--- 622
Db 564 ASDEKCFRPPVTKFKNWSTIYKQGLGVNNEILLTLYTLETQENTYFYQAKTDMYIY 623
QY 623 APGVCVSRPVPVFAFNGSEPEVEGQLGEDNELLPGRELVEPCTANHKKRYFRGADVYV 682
Db 624 KNYEHLKTVPLSSITTLDTIALNFTLLENVDKVIYELTYRDEKRLSNVFDIETMPREYN 683
QY 683 ENYAVYRVPFLAELEVISITFDLNLTVLEDRFLEPVEYTRAEADTGLLDYSEIORNQ 742
Db 684 YYAQRVSGLRKLDLIDLTNRNQFVDAFGSLMDLGLGAVGQTVWNAVSGVATLFSISVTGFI 743
QY 743 LHELRFYDI--DRVVKTDGNMIMRGLANFQOGLGAVGQAVGTVLVGAAGAALSTVSGIA 800
Db 744 NFKNPPGMLMIIVTGVLFATYFLFKTKIVETAPIKMIYPEIDKLKREKGSIEAPI 803
QY 801 SFIANPFGALATGILLVAGLVAFLAYRYISRLSRNPMKALYPTITRAALDDARGATAPG 860
Db 804 SESE-----LER-----IVLAWHIHQONSHMETKTRDKPDS--ILTR-AQNMRLKRS--G 849
QY 861 EEEEFDAALQOAREMIKYMSLVSAVEROEHKAKSKNGGPLLATRLTLALERRAPPE 920
Db 850 YSNLKNAE 857
QY 921 YQQLPMAD 928
```

RESULT 16

```
VGLB_HSVIK
ID_VGLB_HSVIK STANDARD; PRT; 904 AA.
AC P06437;
DT 01-JAN-1988 (Rel. 06, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Glycoprotein B precursor.
```

```
GN OR UL27.
OS Herpes simplex virus (type 1 / strain KOS).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Simplexvirus.
OX NCBI_TaxID=10306;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=84174058; PubMed=6324454;
RA Bzik D.J., Fox B.A., Deluca N.A., Person S.;
RT "The nucleotide sequence of the glycoprotein gene, gB, of herpes
RT simplex virus type 1."
RL Virology 133:301-314(1984).
RN [2]
RP REVISIONS. SEQUENCE FROM N.A.
RX MEDLINE=87071654; PubMed=3024391;
RA Bzik D.J., Debroy C., Fox B.A., Pederson N.E., Person S.;
RT "The nucleotide sequence of the gB glycoprotein gene of HSV-2 and
RT comparison with the corresponding gene of HSV-1."
RL Virology 155:322-333(1986).
RN [3]
RP REVISIONS.
RA Pederson N.E.;
RL Submitted (APR-1987) to the EMBL/GenBank/DBJ databases.
CC -!- SUBUNIT: DIMER, PROBABLY LINKED BY DISULFIDE BONDS.
CC -!- MISCELLANEOUS: THERE ARE SEVEN EXTERNAL GLYCOPROTEINS IN HSV1: GH,
CC GB, GC, GG, GD, GI, AND GE.
CC -!- MISCELLANEOUS: GB IS THE ONLY GLYCOPROTEIN THAT IS KNOWN TO BE
CC REQUIRED FOR VIRAL GROWTH.
CC -!- SIMILARITY: BELONGS TO THE HERPESVIRUSES GLYCOPROTEIN B FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch)
CC -----
DR EMBL; K01760; AAA45774.1; --
DR InterPro; IPR000234; Glycoprot_B.
DR Pfam; PF00606; Glycoprotein_B; 1.
DR ProDom; PD000693; Glycoprot_B; 1.
DR Glycoprotein; Transmembrane; Signal.
KW SIGNAL 1 30
FT CHAIN 31 904 GLYCOPROTEIN B.
FT DOMAIN 31 730 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 731 746 POTENTIAL.
FT TRANSMEM 752 771 POTENTIAL.
FT TRANSMEM 775 795 POTENTIAL.
FT DOMAIN 796 904 CYTOPLASMIC (POTENTIAL).
FT CARBOHYD 87 87 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 141 141 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 398 398 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 430 430 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 489 489 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 674 674 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 904 AA; 100368 MW; B97B7F8DE5FBA299 CRC64;
Query Match 18.7%; Score 849.5; DB 1; Length 904;
Best Local Similarity 27.7%; Pred. No. 1.4e-44;
Matches 245; Conservative 182; Mismatches 358; Indels 101; Gaps 28;
QY 25 LSIAETGVTSPPNTATWSTESP--LTGHYGTHTDSSHGE-RGNENRDSBQNKNIYGSPS 81
Db 59 LGAAPTGDPKPKKKNKPNTPPRPADGNATVAAGHATLREHLRDKAENTDANFY---- 114
QY 82 TFPYRVCSASGVGDVRFQTDHVCPSAD-MVHSEGLLLYKONIIFPMFRVKYRKVVT 140
Db 115 -----VCPPTGATVVQFQPRCPTPRPGQNYTEGIAVFKENIAPYKFKATMYKQVT 169
QY 141 TSTV-----YNGIYSDSITNQHTFFKYSIEPWE--TEKMDTIYOCFNSLRNLNTG 190
Db 170 VSQVWFEGHRSQFMGIFED-----RAPVFEVIDKINAKGVCRTAKYVRNLET 220
```

```

QY 191 TVVDRDDINMTVFLQPVGV-----TPDKRYGSOPELYLEPGWFGWGYRRRTTVC 242
Db 221 TAFHRDDHETDMLKPANAARTSRGWHHTDLKYNPSRVEAF-----HRYGTTVNC 271
QY 243 ELMDFARSNPPDFVATGDTVMSPFWS-GEEDH-ENKMHKPFVSVINNY-KVVD 299
Db 272 IVEEDARSVYPDEFVLATGDFVMSPFYGYREGSHTHTTYAADRFKQVDGFYARDLT 331
QY 300 YQNRCTVPLGKTRIFLDREYVTLSEKHLKMSYCPPLTLKAFYNGIOTESHGSHVFVAN 359
Db 332 TKARATAP--TTRNLTTPKFVANDWVFKRPSVCTMTKWQEVDMLESEYGGSPRFSSD 389
QY 360 DITASFTT-----SKEDMKFNTTYHCLNEBEIKAEIEKKYA-KVNSTHSHKYGLDKLYFK 411
Db 390 AISTFTTNLTLEYPLSRVDLGD-----CIGKDARDAMDRIFARRYNATHIKVGQPIYQ 443
QY 412 TDGGLYLVWQPLQNRLLDAKKNLNNYTSRRSRQAESTTDPMMWMTGNAGGYSSEN 471
Db 444 ANGSGFLIAYQPLSLTLAE---LVVREHLRQSRKPPNPTPPPGAGANASVERIKTTS 499
QY 472 SITVAQOVAYDNLIRINNILEDLSKACREOHRALVNMELSKINPTSVMSMIYNRPV 531
Db 500 SIEFARLQFTYHQRHNDMLGRVAIAWCELQNHLELTMNEARKLNPNATASVTVGRRV 559
QY 532 SAKRIDGVISVNCIIVDQTSVSLHKSRLLSASDEKFSRPPVTFKFMNDSTIYKGLQG 591
Db 560 SARMLGDVMAVSTCVPAADNVIVQNSMR-ISSRPGACYSRPLVSFRYEDQGLVEGQLG 618
QY 592 VNNEILLTTLTLETQENTYFYQAKDMYIYKNEYHLKTVPLSSITLDTFIALNFTLL 651
Db 619 ENNELRLTRDAIEPTCVGHRRYFTFGGVYVPEEYAYSHQLSRADITTVSTFDLNLITML 678
QY 652 ENVDKVELYTRDEKRLSNVPDIETMREYNYAQRVSLGRKLDLSTNRNQF--VDA 709
Db 679 EDHEFVPLEYTRHEIKOSGLLDYTEVQRRLDLRFADI-DTVIHADANAAMPAGUGA 737
QY 710 FGSLMDDLQ-AVGQTVVNAVSGVATLFSISVITGFINFKNPFGGMLIIVIGVLFALYF 768
Db 738 PFEGMGDLGRAVKVMGIVGVW---SAVSGVSSFSNPFALAVGLLVLAGLAAFAFF 793
QY 769 LTKTKYIYETAPIKMIYPEIDK-LK-----EREGKSEIAPISE-BELERIVIAM 815
Db 794 AFYVVMVLQSNPKALYPLTTKELKNPTNPASGEGEEDGFDEAKLAEREMIRYMALV 853
QY 816 HIHQNSHMETKRPDKDSILT-RAQNML---RKESGYNLKNAE 857
Db 854 SAMERTEH---KAKKKGTSAILLSAKVTDMVMKRRNTNYTPVKNK 896

RESULT 17
VCLB_HSV1F STANDARD; PRT; 903 AA.
AC P06436;
DT 01-JAN-1988 (Rel. 06, Created)
DT 01-JAN-1988 (Rel. 06, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Glycoprotein B precursor.
GN GB OR UL27.
OS Herpes simplex virus (type 1 / strain F).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Simplexvirus.
OX NCBI_TaxID=10304;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=85083254; PubMed=2981343;
RA Pellett P.E., Kousoulas K.G., Pereira L., Roizman B.;
RT "Anatomy of the herpes simplex virus 1 strain F glycoprotein B gene:
RT primary sequence and predicted protein structure of the wild type and
RT of monoclonal antibody-resistant mutants.";
RL J. Virol. 53:243-253 (1985).
RN [2]
RP SEQUENCE OF 1-176 FROM N.A.

```

```

RX MEDLINE=88306232; PubMed=2457278;
RA Hammerschmidt W., Conraths F., Mankertz J., Buhk H.-J., Pauli G.,
RA Ludwig H.;
RT "Common epitopes of glycoprotein B map within the major DNA-binding
RT proteins of bovine herpesvirus type 2 (BHV-2) and herpes simplex
RT virus type 1 (HSV-1).";
RL Virology 165:406-418 (1988).
CC -!- SUBUNIT: DIMER, PROBABLY LINKED BY DISULFIDE BONDS.
CC -!- MISCELLANEOUS: THERE ARE SEVEN EXTERNAL GLYCOPROTEINS IN HSV1: GH,
CC GB, GC, GG, GD, GI, AND GE.
CC -!- MISCELLANEOUS: GB IS THE ONLY GLYCOPROTEIN THAT IS KNOWN TO BE
CC REQUIRED FOR VIRAL GROWTH.
CC -!- SIMILARITY: BELONGS TO THE HERPESVIRUSES GLYCOPROTEIN B FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.ebi.ac.uk/announcements/
CC or send an email to license@ebi.ac.uk).
CC
CC EMBL; M14164; AAA45776.1; --
CC EMBL; M21633; AAA45788.1; --
CC PIR; A03750; VGBEB1.
CC InterPro; IPR000234; Glycoprot B.
CC Pfam; PF00606; Glycoprotein B_1.
CC ProDom; PD000693; Glycoprot_B_1.
CC Glycoprotein; Transmembrane; Signal.
FT SIGNAL 1 29
FT CHAIN 30 903 GLYCOPROTEIN B.
FT DOMAIN 31 729 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 730 745 POTENTIAL.
FT TRANSMEM 751 770 POTENTIAL.
FT TRANSMEM 774 794 POTENTIAL.
FT DOMAIN 795 903 CYTOPLASMIC (POTENTIAL).
FT CARBOHYD 86 86 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 140 140 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 397 397 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 429 429 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 488 488 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 673 673 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 903 AA; 100104 MW; 73BCA7813DB35E8 CRC64;

Query Match 18.6%; Score 845.5; DB 1; Length 903;
Best Local Similarity 27.7%; Pred. No. 2.5e-44;
Matches 245; Conservative 179; Mismatches 358; Indels 101; Gaps 28;

QY 28 AETGVTSPPTATWSTESP--LTGHYGHTHDSHGE-RGNNEENRDSSEONKNIYGSPTFP 84
Db 61 APTGDTKPKKKPKPPPPPPAGDNATVAAGHATLREHLRDKAENTDANFY----- 113
QY 85 YRVCSASGVGVFRFQTDHVCPSDAD--MVHSEGLLIYKONIIPFMFRVKRYKVVTTST 143
Db 114 --VCPPTGATVVGFEQPRCPTRPEQGNVTEGIADVFKENIADPYKFKATMYKDVTSQ 171
QY 144 V-----YNGIYSDISITNQHTFYKSIQWE--TEKMDTIYQCFNSLRNLGNTLYV 193
Db 172 VWFGRYSQFMGIFED-----RAPVPFEEVIDKINAKGVCRTAKYVRNNLETTAF 222
QY 194 DRDDINMTVFLQPVGV-----TPDKRYGSOPELYLEPGWFGWGYRRRTTVCNCLM 245
Db 223 HRDDHETDMLKPANAARTSRGWHHTDLKYNPSRVEAF-----HRYGTTVNCIVE 273
QY 246 DMFARSNPPDFVATGDTVMSPFWS-GEEDH-ENKMHKPFVSVINNY-KVVDYQN 302
Db 274 EVDARSVPYDEFLATGDFVMSPFYGYREGSHTHTTYAADRFKQVDGFYARDLTAKA 333
QY 303 RGTVPCLKTRIFLDREYVTLSEKHLKMSYCPPLTLKAFYNGIQTESHGSHYHVANDIT 362
Db 334 RATAP--TTRNLTTPKFVANDWVFKRPSVCTMTKWQEVDMLESEYGGSPRFSSDAIS 391
QY 363 ASPTT-----SKEDMKFNTTYHCLNEBEIKAEIEKKYA-KVNSTHSHKYGLDKLYFKTDG 414

```

```

Db 392 TTTTNTLTPLSRDLGD-----CIGKDARDMDRIFARRYNATHIKVGQOYYLANG 445
QY 415 GLYLVMQPLQLNRLDLAKNKLNNETYSRRSRQAESTTDPMMEMTNGAGGEYSSENSIT 474
Db 446 GFLLAYOPLLSNTLAE-----LYVREHLREQSRKPPNTPPPPCASANASVERIKTTSSIE 501
QY 475 VAQOYAYDNLRIINILDLKAMCREQHRAALVWNLKSKINPTSVMSMIYNRPVSAS 534
Db 502 FARLOQTYNHQIRHVNDMLGRVAIAWCELQNHLETLWNEARKLPNPAIASATVGRVSAR 561
QY 535 RIGDVTLSVNCIIVDQTSVLSHLKSLRLSDEKCFSPRPVPTFKWMDSTIYKQGLGVNN 594
Db 562 MLGDVWAVSTCVPAADNVIVQNSMR--ISSRPGACYSRPLVSFRYEDQGLVEGQLENN 620
QY 595 EILLTTLTTCENTETYYFQAQTDVMIYKYNHKLKTPVLSSITTLDTFTIALNFTLENV 654
Db 621 ELRLTRDAIEPCTVGHRRYTFGGYVYFEEYAYSHQLSRADITTVSTFTDLNITWLEDH 680
QY 655 DFKVIELYTRDEKRLSNVFDIETWFRYNYAORVGLRKLDDLSTNRQF--VDAFGS 712
Db 681 EFPVLEYVTRHEIKDSGLDYTEVQRNQLHDLRFADI-DTVIHADANAAMFAGLGAFFE 739
QY 713 LMDDLG-AVGOTVNAVSGVATLFSSTVTCFINFKNPGCGMLMIIVIGVLEAIVFLTK 771
Db 740 GMDGLGRAVKVWVGIVGVV-----SAVGSVFSMNPFGALAVGLLVLAGLAAAFAPR 795
QY 772 KTKIYETAPIKMIYPEIDK-LK-----EREKGSBIAPISE-EELERIVLAMHIIH 818
Db 796 YVMRLQSNPMKALYPLTTKELKNPTNPDASGEEGGDFDEAKLAAREMIRYALVSAM 855
QY 819 QQNSHMETKTRKDPKDSILT-RAQNML---RKSGYSNLKNAE 857
Db 856 ERTEH---KAKKGTSALLSAKVTDVMVRKRNTNTYQVFNKD 895

RESULT 18
VGLB HSV2H STANDARD; PRT; 904 AA.
AC P08666; P89450;
DT 01-JAN-1988 (Rel. 06, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Glycoprotein B precursor.
GN GB OR UL27.
OS Herpes simplex virus (type 2 / strain HG52).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Simplexvirus.
OX NCBI_TaxID=10315;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=87071654; PubMed=3024391;
RA Bzik D.J., Debroy C., Fox B.A., Pederson N.E., Person S.;
RT "The nucleotide sequence of the gB glycoprotein gene of HSV-2 and
RT comparison with the corresponding gene of HSV-1.";
RL Virology 155:322-333(1986).
RN [2]
RP SEQUENCE FROM N.A.
RA Dolan A.;
RL Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.
CC -1- SUBUNIT: DIMER, PROBABLY LINKED BY DISULFIDE BONDS.
CC -1- MISCELLANEOUS: THERE ARE SEVEN EXTERNAL GLYCOPROTEINS IN HSV1 AND
CC 2: GH, GB, GC, GD, GI, AND GE.
CC -1- MISCELLANEOUS: GB IS THE ONLY GLYCOPROTEIN THAT IS KNOWN TO BE
CC REQUIRED FOR VIRAL GROWTH.
CC -1- SIMILARITY: BELONGS TO THE HERPESVIRUSES GLYCOPROTEIN B FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/

```

```

or send an email to license@isb-sib.ch).
CC
CC EMBL; M14923; AAA66440.1; -.
DR EMBL; Z86099; CAB06752.1; -.
DR InterPro; IPR000234; Glycoprot B.
DR Pfam; PF00606; Glycoprotein B; 1.
DR ProDom; PD000693; Glycoprot_B; 1.
KW Glycoprotein; Transmembrane; Signal.
FT SIGNAL 1 22
FT CHAIN 23 904
FT DOMAIN 23 727
FT TRANSMEM 728 743
FT TRANSMEM 749 768
FT TRANSMEM 772 792
FT DOMAIN 793 904
FT CARBOHYD 82 82
FT CARBOHYD 136 136
FT CARBOHYD 393 393
FT CARBOHYD 425 425
FT CARBOHYD 486 486
FT CARBOHYD 671 671
FT CONFLICT 92 92
FT CONFLICT 198 198
FT CONFLICT 308 308
FT CONFLICT 438 438
FT CONFLICT 568 568
FT CONFLICT 619 620
FT CONFLICT 636 636
FT SEQUENCE 904 AA; 100217 MW; AB050A3AFB4F1066 CRC64;
Query Match 18.6%; Score 844; DB 1; Length 904;
Best Local Similarity 28.4%; Pred No. 3.1e-44;
Matches 247; Conservative 164; Mismatches 366; Indels 92; Gaps 27;
QY 33 TSPPTATWSTESPLTCHYTHDSSHGERRGNENRDESEQKNKIYGSPTFPYRVCASG 92
Db 76 TTPPD---ANATVAAGHATLR--AHUREIKVENADAQ-----FVCPPT 115
QY 93 VGDVFRFQDTHVCPDASD-MVHSEGLLLYKQNIIFPMFRVRYKRVKVVITSTV----- 144
Db 116 GATVQFEQPRRCPTREPGQNYTEGIAVFKENIAPYKFKATMYKDVTSQVWFGHYS 175
QY 145 -YNGIYSDSITNQHTFYKSIPEWE--TEKMDTIYQCFNSLRLNTGNNLLTYVDRDDINMT 201
Db 176 QFMGIFD-----RAPVPEEVIDKINTKVCSTAKYVNRNMETFAFRDHDHTD 226
QY 202 VFLQPVGDVTPDVKRYGSOPELYLEPGFWGYSYRRTTVNCELMDMFARSNPFPFVTA 261
Db 227 MELKPAK-VATRTSRGWHHTDLKYNPSRVEAFHRYGTIVNCIVEVDARSVPVDFVLA 285
QY 262 TGDVTMSFPWS-GEDDH-BNKMHEKPFVSVVINNY-KVVDYQNRGTVPGLKTRIFLDRE 318
Db 286 TGFVYMSFPFYGYREGSHTSHTSYAADRFKQVDGFYARDLTTKARATSP--TTRNLLTTP 343
QY 319 EYTLSEKHLKMSYCPLTLKAFYNGIOTEHSGSYHFVANDITASFTT-----SKED 371
Db 344 KFTVAMDVWPKRPVACTMTKNQDEVDMLEAERYGCGSFRFSSDAISTFTTNLTYSLSRVD 403
QY 372 MKEFNTHYCLNEEIKAEIEKYYA-KVNSTHSKYGLKFKPTDGLLYVMQPLQLNRLD 430
Db 404 LGD-----CIGRDAREADRMFARKYNATHIKVGQOYYLATGGFLIAYOPLLSNTLAE 457
QY 431 AKNKLANNETYSR-RSRQAEESTDPMEM-TGNAGAGEYSSSENSITVAQOYAYDNLRI 488
Db 458 ----LVVREYMRQDQKPRNATPALEAPASNAVERIKTSSIFARLQFTYNNHQRH 513
QY 489 INNILEDLSKAWCREQHRAALVWNLKSKINPTSVMSMIYNRPVSASAKRIGVISVNCIV 548
Db 514 VNDMLGRIAVAWCELQNHLETLWNEARKLPNPAIASATVGRVSARMLGDVMAVSTCVPV 573
QY 549 DQTSVLSHLKSLRLSDEKCFSPRPVPTFKWMDSTIYKQGLGVNNELITTTVLETCQE 608
Db 574 APDNVIVQNSMR-VSSRPGTCYSRPLVSFRYEDQGLIEGOLGNNELRLTRDALEPCTV 632

```



```

ID VGLB_HSV23 STANDARD; PRT; 904 AA.
AC P06763;
DT 01-JAN-1988 (Rel. 06, Created)
DT 01-JAN-1988 (Rel. 06, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Glycoprotein B precursor.
GN GB OR UL27 OR GB2.
OS Herpes simplex virus (type 2 / strain 333).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Simplexvirus.
OX NCBI_TaxID=10313;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=87112925; PubMed=3027364;
RA Stuve L.L., Brown-Shimer S., Pacht C., Najarian R., Dina D.,
RA Burke R.L.;
RT "Structure and expression of the herpes simplex virus type 2
RL J. Virol. 61:326-335(1987).
CC -1- SUBUNIT: DIMER, PROBABLY LINKED BY DISULFIDE BONDS.
CC -1- MISCELLANEOUS: THERE ARE SEVEN EXTERNAL GLYCOPROTEINS IN HSV1 AND
CC 2: GH, GC, GD, GI, AND GE.
CC -1- MISCELLANEOUS: GB IS THE ONLY GLYCOPROTEIN THAT IS KNOWN TO BE
CC REQUIRED FOR VIRAL GROWTH.
CC -1- SIMILARITY: BELONGS TO THE HERPESVIRUSES GLYCOPROTEIN B FAMILY.
CC
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC -----
DR EMBL; M15118; AAA45837.1; -.
DR PIR; A26790; VGBEB2.
DR InterPro; IPR000234; Glycoprot B.
DR Pfam; PF00686; Glycoprotein B_1.
DR ProDom; PD000693; Glycoprot B; 1.
KW Glycoprotein; Transmembrane; Signal.
FT SIGNAL 1 22 POTENTIAL.
FT CHAIN 23 904 GLYCOPROTEIN B.
FT DOMAIN 23 727 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 728 743 POTENTIAL.
FT TRANSMEM 749 768 POTENTIAL.
FT TRANSMEM 772 792 POTENTIAL.
FT DOMAIN 793 904 CYTOPLASMIC (POTENTIAL).
FT CARBOHYD 82 82 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 136 136 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 393 393 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 425 425 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 486 486 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 671 671 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 904 AA; 100186 MW; A8B36F74FDBC839 CRC64;

Query Match
Best Local Similarity 28.3%; Pred. No. 6.4e-4;
Matches 246; Conservative 164; Mismatches 367; Indels 92; Gaps 27;

QY 33 TSPNTATWSTESPLTGHYTHDSHGCGNENRDSSEQNKIYGPSPTFPYRVCSASG 92
DB 76 TTPPD-----ANATVAAGHATLR--AHLREIKVENADAQ-----FYVCPPT 115
QY 93 VGDVRFQTDHVCDSAD-MVHSEGLLIYKQNIIPFMFRVRYKRVKVVTTSTV----- 144
DB 116 GATVVOEQPRRCPTREGQNYTEGIAVFEKENIAPYKATMYKDVTSVQVWFGRHYS 175
QY 145 -YNGIYSDSTNQHTFYKSTPEWE--TEKMDTIYQCNSRLNTGNNLLTYVDRDDINMT 201
DB 176 QFMGIFED-----RAPVPEEVIDKJNAKGVCRSTAKYVRNNMETTAFHRDHDHTD 226
QY 202 VFQIPVDGVTPDKRYGSOPELYLEPGFWGVSRRRTTVNCELMDMFARNPFPDFEFTA 261

```

CC -1- SIMILARITY: BELONGS TO THE HERPESVIRUSES GLYCOPROTEIN B FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

DR EMBL: K03541; AAA45778.1;
 DR InterPro: IPR000234; Glycoprot B.
 DR Pfam: PF00606; Glycoprotein B; 1.
 DR ProDom: PD000693; Glycoprot B; 1.
 KW Glycoprotein; Transmembrane; Signal.
 FT SIGNAL 1 30
 FT CHAIN 31 904
 FT DOMAIN 31 730
 FT TRANSMEM 731 746
 FT TRANSMEM 752 772
 FT TRANSMEM 775 795
 FT DOMAIN 796 904
 FT CARBOHYD 87 87
 FT CARBOHYD 141 141
 FT CARBOHYD 398 398
 FT CARBOHYD 430 430
 FT CARBOHYD 489 489
 FT CARBOHYD 674 674
 SQ SEQUENCE 904 AA; 100115 MW; 782581DC830A626F CRC64;

Query Match 18.4%; Score 836.5; DB 1; Length 904;
 Best Local Similarity 27.5%; Pred. No. 9.1e-44;
 Matches 244; Conservative 181; Mismatches 361; Indels 101; Gaps 28;

QY 24 SLSTAETGVTPPNTATWSTSP--LTGHYTHDSSHGE-RGNENRDSSEQNKNYICSP 80
 DB ALGAAPTGDPPKPKKPKNPPRPPAGDNATVAAGHATLREHLRDIKAENTDANFY--- 114
 QY 81 STFPYRVCSAGVGDFRFTQDTHVCPDASD-MVHSEGLILLYKONIIPFMRVRYKRV 139
 DB -----VCPPTGATVGFQPRRCPTRPGEQNYTEGIAVFKENIAYKFKATMYKDV 168
 QY 140 TTSTV-----YNGIYSDSITNHTFYKSIEMPE--TEKMDTIYQCFNSLRNTG 189
 DB TVSQWVFGHRYSQMGIFED-----RAPVPEEVIDKINAGVCRSTAKYVRNLE 219
 QY 190 LTYVDRDDINTVFLQPDGV-----TPDKYGSQPELYLEPGFWGSGYRRRTTVN 241
 DB TFAHRDDHETDMLKPNAAATRTSRGWHTTDLKYNPSRVEAF-----HRYGTTW 270
 QY 242 CELDMFARSNPPDFVFTATGDTVMSPFWS-GEDDH-ENKMHEKPFVSVINNY-KV 298
 DB CIVEVDARSVPYDFVLATGDFVYMGFFVGYREGSHTHTSYAADRKFQVDGFYARDL 330
 QY 299 DYQNRGTVPGLKTRIFLDREYTLISWEKHLKNMSYCLPLTWKAFYNGIQTHSGSYH 358
 DB TTKARATAP--TTRNLLTPKFTVAWDVVPKPSVCTWKQEVDEMRLSEYSGSFRESS 388
 QY 359 NDITASFT-----SKEDMKEFTYTHCLNEEIKABIEKYYA-KVNSTHSKYGDLKY 410
 DB DAISTFTFTNTTEYPLSRVDLGD-----CIGKDARDAMDRIFFARRYNATHIKVG 442
 QY 411 KTDGGLVWOPLOINRLLDAKNLNNETYRRSRROAESTTDPMMWNTGNGAGEYSSE 470
 DB LANGGFLAYQPLSNTLAE-----LYVREHLREQSRKPPNTPPPPGASANASVERIK 498
 QY 471 NSITVAQVQYAYDNIIRINNIEDLSKAWCREQHRALVWNLSEKINPTSVMSMYNRP 530
 DB SSIEFARLQFTYNHQHRVNDMLGRVATWCELQNHLETLNNEARKLPNPAIASATVGR 558
 QY 531 VSAKRIGDVISVNCIVVDQTSVLHKSRLLSASDEKCFRPPVPTFKFMDNSTIYKQL 590
 DB VSARMLGDVMAVSTCVPAADNVIVQNSMR-ISSRPGACYSRPLVSPRYEQDGLVEQ 617

QY 591 GWNNEILLTITTYLETQENTYFYFOAKTDMYIKNYEHLKTVPLSSITTLDTFIALNFTL 650
 DB GENNELRLTRAIDPCTVGHRYTFTGGGYVYFESAYSHQLSRADITTVSTFIDLNTM 677
 QY 651 LENVDKVIETYLRDEKRLSNVFIETMFRFNYNYYAQRVSLRDLDLSTNRNOF--VD 708
 DB LEDHEFVPLEYVTRHEIKDGLDYTEVQRNQLHDLRFADI-DTVIHADANAAMFAGLG 736
 QY 709 AFGSLMDDLG-AVGQTVVNAVSGVATLFPSSIVTGFINFIPKPRGMLMIIVIGVLPAY 767
 DB AFPEGMDLGRAVGKVMGIVGGVV-----SAVGSVSSPMSPFGALAVGLLVLAGLAAAF 792
 QY 768 FLTKTKIYETAPIKMIYPEIDK-LK-----EREKSEIAPISE-EELERIVLA 814
 DB FAFRYVMRLQSNPKALYPLTTKELKNPTNPDASGEGEEDGDFDEAKLAAREMRYMAL 852
 QY 815 MHIHQNSHMETKTRKOPKDSILT-RAQNML---RKRSGYSNLKNAE 857
 DB VSAMERTEH---KAKKGTSAALLSAKVTDMVMRKRRTNTYQVFNKD 896
 RESULT 22
 VGLB ILTVS STANDARD; PRT; 883 AA.
 ID VGLB ILTVS
 AC P27415;
 DT 01-AUG-1992 (Rel. 23, Created)
 DT 01-AUG-1992 (Rel. 23, Last sequence update)
 DT 16-Oct-2001 (Rel. 40, Last annotation update)
 DE Glycoprotein B precursor.
 GN GB.
 OS Infectious laryngotracheitis virus (strain SA-2) (ILTV).
 CC Viruses: dsDNA viruses, no RNA stage; Herpesviridae;
 CC Alphaherpesvirinae; Infectious laryngotracheitis-like viruses.
 CC NCBI_TaxID=10343;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=91335774; PubMed=1840710;
 RA Kongsuwan K., Pridaux C.T., Johnson M.A., Sheppard M., Fahey K.J.;
 RT "Nucleotide sequence of the gene encoding infectious
 RT laryngotracheitis virus glycoprotein B";
 RL Virology 184:404-410(1991).
 CC -1- SIMILARITY: BELONGS TO THE HERPESVIRUSES GLYCOPROTEIN B FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: M64927; AAA88009.1; ALT SEQ.
 DR PIR: A40567; VGBEIS.
 DR InterPro: IPR000234; Glycoprot B.
 DR Pfam: PF00606; Glycoprotein B; 1.
 DR ProDom: PD000693; Glycoprot B; 1.
 KW Glycoprotein; Transmembrane; Signal.
 FT SIGNAL 1 26
 FT CHAIN 27 883
 FT DOMAIN 27 724
 FT TRANSMEM 725 744
 FT TRANSMEM 751 771
 FT DOMAIN 772 883
 FT CARBOHYD 102 102
 FT CARBOHYD 121 121
 FT CARBOHYD 211 211
 FT CARBOHYD 262 262
 FT CARBOHYD 360 360
 FT CARBOHYD 579 579
 FT CARBOHYD 635 635
 FT CARBOHYD 649 649
 SQ SEQUENCE 883 AA; 100142 MW; B414083346C55014 CRC64;

```
Query Match      18.4%; Score 836; DB 1; Length 883;
Best Local Similarity 27.9%; Pred. No. 9.4e-44;
Matches 222; Conservative 158; Mismatches 308; Indels 108; Gaps 23;

OY 57 SHGERGNENRD--SEQNKNY-----GSPSTFPYR-----VCSASGVGDVFRFDHV 104
DB 34 SHGIGWNSPHDTASMDVGKISFSEALGSGAPKEPQIRNRIFACSSPTGASVLAQPRH 93
OY 105 CPDASDVH--SEGILLLYKQNIIPMFVRKYRKVTTSTVYNGIYSDSTNQHTFYKSI 163
DB 94 CHRHADSTNMTAGIAVFKQNIAPVFNVTLYKHITTTWALFSRPOITNEYVTRVPI 153
OY 164 EPWETKMDITIOCFNSLRNTGNNLTYYDVRDINMTVLOP-----VDGVTPDV 214
DB 154 DYHEIVRIDRSGECSSKATYHKNFMPFEAVDNDAEKKLPVLPBLLRSTVSKAPHTTNT 213
OY 215 KRYGQPELYLEPGWFGWYRRRTTVCNCELMDFARSNPPDFEVTATGDTVMSPFWSG 274
DB 214 KRHQTL-----GYRTSTSDVCVVEYLQARSVPYDYEGMATGDTVELSPFYT- 260
OY 275 EDDHENQKHPWFSVINNYK---VVDYQNR-----GTVPLGKTRIFLDREYTLSEKH 327
DB 261 -----KNITGPRHRSYRDYRFLFIANYQVRDLETGQIRPPKKRNFELTDEQFTIGWAM 314
OY 328 LKMSYCPCLTLKAFNGIQTEHSGSYHFVANDITASFTTSKEDMKFNNT-----YHCL- 382
DB 315 EEKESVCTLKGWIEVPAVRVSNYSYHFLSKDWTMTFFSGKQ---PFIISRLHLAECVP 371
OY 383 ---NEEKAEIEKKYAKVNSTHSHYKGLYKPKTDGGLYLVNQPLIQNRLLD---AKNKLN 436
DB 372 TIASEADIGIPARKY---SSTHVRSGLIEYVYLGSGGLIAFQKLSHGLAEMLYEEAQRQ 428
OY 437 NETYSRRSRQAESTTDPMMEMTNGAGGEYSSENSITVAQVQVAYDNLIRINILEDL 496
DB 429 NHPGRERRRQAQRRTASLQ--SGPQGDRIITHTSSATFAMLOPAYDKIOAHVNNELIGNL 486
OY 497 SKACRQHQRAALVWNLKINPTSVMSYMLNRPVSAKRIGDVLISVNSCIVVDOTSVSLH 556
DB 487 LEAWCEIQNLQIVVHWEKLNPNLSMTSLFGQVPSARLLGDIVAVASKIEIPIENTRMQ 546
OY 557 KSLRLLSASDEKFSRPPVTFKFM-----NDSTIYKQGLGVNNEILLTTTLYL 603
DB 547 DSMR-VPGDPTMCTYRVLIFRYSSSPESQFSANSTENHNLGILQGLGEHNEILQGRNLI 605
OY 604 ETCOENTYFPQAKTMYIKVNEHLKTVPLSSITTLDTFIALNFTLLENVDFKVIELYT 663
DB 606 EPCMINHRRYFLGENTLYLLEDYTFVRQVNASEIEEVSFTINLNTATLEDLDFVPVEVYT 665
OY 664 RDEKRLSNVFDIETMFREYNVYQAR-----VSGLRKDLDDLSTNRNQFVDAFGSLMDD 716
DB 666 REELRDTGTYNDVVRQYQNIYKRFIDITVIRGDRGDAI-----PRAIADFNGNTLGE 720
OY 717 LG-AVGQTVVNAVSGVATLFSSIVTGFPIFKNPFQGM-----LMIITVWIGVLFAIYFLT 770
DB 721 VGRALGTVMVTAATAAVI-----STVSGIASFLSNPFAALGIAVAVVSIILGLLAFKYMVN 776
OY 771 KTKIYETAIKMIYP 786
DB 777 LKSN-----PVQVILFP 787

RESULT 23
VGLB_HSVB
ID VGLB_HSVB STANDARD; PRT; 980 AA.
AC P28922;
DT 01-DEC-1992 (Rel. 24, Created)
DT 01-DEC-1992 (Rel. 24, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Glycoprotein B precursor (Glycoprotein 14).
GN GB OR GP14 OR 33.
OS Equine herpesvirus type 1 (strain Ab4p) (EHV-1), and
OS Equine herpesvirus type 1 (strain Kentucky D) (EHV-1).
```

```
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Varicellovirus.
OX NCBI_TaxID=31520, 10330;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=AB4P;
RX MEDLINE=92295566; PubMed=1318606;
RA Telford E.A.R., Watson M.S., McBride K., Davison A.J.;
RT "The DNA sequence of equine herpesvirus-1.";
RL Virology 189:304-316(1992).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Kentucky D;
RX MEDLINE=90219246; PubMed=2157895;
RA Guo P.X., Goebel S., Ferkus M.E., Taylor J., Norton E., Allen G.,
RA Languet B., Demetree P., Paoletti E.;
RT "Coexpression by vaccinia virus recombinants of equine herpesvirus 1
RT Glycoproteins gp13 and gp14 results in potentiated immunity.";
RL J. Virol. 64:2399-2406(1990).
CC - SIMILARITY: BELONGS TO THE HERPESVIRUSES GLYCOPROTEIN B FAMILY.
CC
CC THIS SWISS-PROT entry is copyrighted. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; M86664; AAB02468.1; -
CC EMBL; M35145; AAA46067.1; -
CC PIR; G36798; VGBEC6.
CC InterPro; IPR000234; Glycoprot_B.
CC Pfam; PF00606; Glycoprotein_B; 1.
CC ProDom; PD00693; Glycoprot_B; 1.
CC Glycoprotein; Transmembrane; Signal.
CC SIGNAL 1 86 POTENTIAL.
CC CHAIN 87 980 GLYCOPROTEIN B.
CC DOMAIN 87 852 EXTRACELLULAR (POTENTIAL).
CC TRANSMEM 853 870 POTENTIAL.
CC DOMAIN 871 980 CYTOPLASMIC (POTENTIAL).
CC CARBOHYD 165 165 N-LINKED (GLCNAC. ) (POTENTIAL).
CC CARBOHYD 275 275 N-LINKED (GLCNAC. ) (POTENTIAL).
CC CARBOHYD 380 380 N-LINKED (GLCNAC. ) (POTENTIAL).
CC CARBOHYD 423 423 N-LINKED (GLCNAC. ) (POTENTIAL).
CC CARBOHYD 497 497 N-LINKED (GLCNAC. ) (POTENTIAL).
CC CARBOHYD 514 514 N-LINKED (GLCNAC. ) (POTENTIAL).
CC CARBOHYD 515 515 N-LINKED (GLCNAC. ) (POTENTIAL).
CC CARBOHYD 560 560 N-LINKED (GLCNAC. ) (POTENTIAL).
CC CARBOHYD 727 727 N-LINKED (GLCNAC. ) (POTENTIAL).
CC CARBOHYD 749 749 N-LINKED (GLCNAC. ) (POTENTIAL).
CC SEQUENCE 980 AA; 109805 MW; A6DDFA8CA5550FF5 CRC64;

Query Match      18.4%; Score 835; DB 1; Length 980;
Best Local Similarity 26.9%; Pred. No. 1.3e-43;
Matches 250; Conservative 161; Mismatches 383; Indels 134; Gaps 31;

OY 32 VTSPNTATWSTESPLTGHYTHDSHGERGNENNRDSEQNKNYGSPTFPYRVC-SA 90
DB 83 VRAVPTTSPPTSTPTS--MSTH--SHG-----TVDPDLLTPTDPLRLAVRE 127
OY 91 SGV-----GD-----VRFQTDHVCDDAS--DMVHSEGLILLYKQNIIPMFVRVK 134
DB 128 SGILAEDEGDFYTCPPPTGTVVRIEPPRTCPKPDGLGRNFTGIAVFKNIAPYKFRANV 187
OY 135 YRKVVTTSTVYNGIYSDSTNQHTFYKSIETPTEK-----MDTIYOCFNS---LRLNTGG 187
DB 188 YKDIIVTVRWKGYSHTSLSDR---YNDRVPSVSEIFGLIDSKGKCSKAEYL- 240
OY 188 NLTYYV---DRDINMTVFLQPDVGVTPDKVRYGSGQPELYLEPGWF-WGYSRRRTTVNCE 243
DB 241 NIMHAYHDEDEVELD--LVPSKATPGARAWQTTNDTTTSYVGMMPWRHY-TSTSVNCI 297
```


Db 500 RHVNDMLGRITAWCEQLQNRRLTWNERRLNPCALASATVGRVRSARMLGDVNAVSTCV 559
 QY 547 VVQTSVSLHKSRLLSASDEKCFSPPTVFKFNDSTIYKGLGVNNEILLTTTILETC 606
 Db 560 PVAPDNVIMONSIG-VAARPGTCYSRPLVSFRYEADGPLVEGQGDNEIRLDRLEPC 618
 QY 607 QENTEYFQAKTDMYIKYNEHLKTVPLSSITITLDTFIALNFTLLENVDFKVIETRDE 666
 Db 619 TVGHRRTFTGAGVYVEEYAYSHQLGRADTVTVSTFIFINLNTWLEDFEVPLEVTRQE 678
 QY 667 KRISNVFDIETMFREYNYYAQRVSGRLKOLLIDSTNRNQVDFAGSLMDLGAAGVQTVN 726
 Db 679 IKDSGLLDYEVQRNQLHALRFADI-DTVIKADAAHALFAGLY-SFFEGLGDVGRAVGK 736
 QY 727 AVSGVATLFSIVTGFNFNFPGGMLIIVIGVLFALYFLTKTKIYETAPIKIYIP 786
 Db 737 VVNGIVGGVVSAYSGVSSFLSNPFGALVGLLVLGALAAFFAFRYVMRLQRNPKALYP 796
 QY 787 EIDKLKERECKSEIAPIS-----EELERIVLA-----MHIHQONSHMETKTRKOP 832
 Db 797 ----LTKELKSDGAPLAGGEGDEGAEFDDEAKLAQAREMIRYMALVSAMERTHEKARKG 852
 QY 833 KDSIL-----TRAQNMLRKRGYGNLKNASVEM 861
 Db 853 TSALLSAKVTDAVMRKARPRYSPLRDTDEEL 885

RESULT 25
 VGLB_VZVD STANDARD; PRT; 868 AA.
 AC P09257;
 DT 01-MAR-1989 (Rel. 10, Created)
 DT 01-MAR-1989 (Rel. 10, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Glycoprotein B precursor (Glycoprotein II).
 GN 31.
 OS Varicella-zoster virus (strain Dumas) (VZV).
 OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
 OC Alphaherpesvirinae; Varicellovirus.
 OX NCBI_TaxID=10338;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=86306657; PubMed=3018124;
 RA Davison A.J., Scott J.E.;
 RT "The complete DNA sequence of varicella-zoster virus.";
 RL J. Gen. Virol. 67:1759-1816(1986).
 CC -/- SIMILARITY: BELONGS TO THE HERPESVIRUSES GLYCOPROTEIN B FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/)
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR ENBL; X04370; CAA27914.1; -;
 DR PIR; E27214; VGBE31.
 DR InterPro; IPR00234; Glycoprot B.
 DR Pfam; PF00606; Glycoprotein_B_1.
 DR ProDom; PD000693; Glycoprot_B; 1.
 KW Glycoprotein; Transmembrane; Signal.
 FT SIGNAL 1 ?
 FT CHAIN ? 868 GLYCOPROTEIN B.
 FT DOMAIN ? 679 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 680 695 POTENTIAL.
 FT TRANSMEM 701 720 POTENTIAL.
 FT TRANSMEM 724 744 POTENTIAL.
 FT DOMAIN 745 868 CYTOPLASMIC (POTENTIAL).
 FT CARBOHYD 84 84 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 194 194 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 372 372 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 440 440 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 557 557 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 623 623 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 868 AA; 98066 MW; 5888E1346F4FF302 CRC64;
 Query Match 18.3%; Score 832; DB 1; Length 868;
 Best Local Similarity 26.7%; Pred. No. 1.6e-43;
 Matches 236; Conservative 168; Mismatches 365; Indels 116; Gaps 27;
 QY 40 TWSTESPLTGHVTHDSSHGGRNNENRDSSEQNKNIIYGSPTFFPYRVCSASGVDFRF 99
 Db 23 TOSDETITRAHLDGSD-----EIRAIHKSODAETKPT---FYVCPPTGSTIVRL 70
 QY 100 QTDHVCDAADMVH-----SEGILLYIKONIIPFMRVRKRVKVVTTSTVNGIYSDSIT 154
 Db 71 ETRICPD---YHLGKNFTGIAVYKENIAAYKEKATVYKDVIVSTAWAGSSYTOIT 126
 QY 155 NQHTFYKSIPEWE-TEKMDTIYQCFNSRLNTGGNLLTYV-----DRDDINMTV 202
 Db 127 NRYADRVPIPVSEITDTIDKFGK-----SSKATVVRNNHKVFAFNEDKNQDMPL 177
 QY 203 FLOPVDGVTDPVKRYGSOPELYLEPCWFNGSYRRRTTVNCELMDMFARSNPPDFVTAT 262
 Db 178 IASKYNSV--GSKAMHTTNDTYMVAG--TPCTYRTGTSVNCIIEVEARSIFPYDSFGLST 234
 QY 263 GOTVEMSPWSGED-----DHEN-----KMHEKWFVSVINNYKVVDYQNRGTVPGLKTRIF 314
 Db 235 GLIIMSFPFGLRDGAYREHSNYAMDRFHQ-----PEGYRQRLDITRALLE-PAARNF 286
 QY 315 LDREYTLSEKHLKNMSYCLPLTKAFYNGIQTEHSGSYHFVANDITASFTTSKEDMKE 374
 Db 287 LVTPHLTVGNWPKPKRTEVCSLVKQREVEDVRDEYAHNFRFTMTKLTSTTFIS---ETNE 343
 QY 375 FNTTY-----HCLNEEIKABIEKKY-AKVNSTHISKYGDLYKFTDGGGLYVWQBLIONRL 429
 Db 344 FNLNQIHLSCQKKEARAIINRIYTRYNSSHVRTGDIQTYLARGGVFVVFQPLLNSLA 403
 QY 430 ----DAKNKLNNETYRRSRROAESTTDPMMEMTGNAGGAGESSENSITVAQVQAYDN 484
 Db 404 RLYQLQELVRENTNHSPOKHPTRNRSRVPVELRANRT---ITTSSEVFAMLQTYD 460
 QY 485 LRIRINNILEDLSKAWCREQHRALVWNELSKINPTSVMSIYNRPVSAKRIGDVISVN 544
 Db 461 IQEHVNEMLARISSSWCQLQNRERALSGLFPINPSALASTILDQRVKARILGDVISVN 520
 QY 545 CIWV-DQTSVSLHKSRLLSASDEKCFSPPTVFKFNDSTIYKGLGVNNEILLTTTIL 603
 Db 521 CPGLGSDTRIIILQNSMR-VSGSTTRCYRPLSIVSLNGSGTVEGQLGTDNEILMSRDL 579
 QY 604 ETCOENTEYFQAKTDMYIKYNEHLKTVPLSSITITLDTFIALNFTLLENVDFKVIET 663
 Db 580 EPCVANHKRYFLFGHHYVYEDYRYVRETAHVDVGMISTYVDNLNLTLLKDRFEMPLOVYT 639
 QY 664 RDEKLSNVFDIETMFREYNYYAQRVSGRLKOL-LDLSTNRNQVDFAGSLMDLGAAGVQ 722
 Db 640 RDELRTDGLDYSEIQRRNQMSHLSRFYDIDKVVQYDSGT---AIMQGMQAFQFGLGTAGQ 696
 QY 723 TVNVAVSGVATLFSIVTGFNFIPKPFPGMLMIIVIGVLFALYFLTKTKIYETAPIK 782
 Db 697 AVGHVVLGATGALLSTVHGFTTFLSNPFGALAVGLLVLGALVAAFFAYRYVLKLTSPMK 756
 QY 783 MLYPEIDK-LKE-REGKSEIA-----PISE-----EELERIV 812
 Db 757 ALYPLTTKGLKQLPEGMDPFAEKPNATDTPIEIGDSQNTSPSVNSGFDPKFREAEQMI 816
 QY 813 LAMHIHQONSHMETKTRKDPKOSILTRAQ---NMLKRSGYGNL 854
 Db 817 KYMTLSAAERQESKARKKNTKSAULTSRLTGLALNRGRYSRVR 861
 RESULT 26
 VGLB_ILVT
 ID VGLB_ILVT
 AC P24904; STANDARD; PRT; 883 AA.

DT 01-MAR-1992 (Rel. 21, Created)
 DT 01-MAR-1992 (Rel. 21, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Glycoprotein B precursor.
 GN GB.
 OS Infectious laryngotracheitis virus (strain Thorne V882) (ILTV).
 OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
 OC Alphaherpesvirinae; Infectious laryngotracheitis-like viruses.
 RN NCBI_TaxID=10344;
 RX SEQUENCE FROM N.A.
 RP MEDLINE=91132136; PubMed=1847176;
 RA Griffin A.M.;
 RT "The nucleotide sequence of the glycoprotein gB gene of infectious laryngotracheitis virus: analysis and evolutionary relationship to the homologous gene from other herpesviruses.";
 RL J. Gen. Virol. 72:393-398 (1991).
 CC -!- SIMILARITY: BELONGS TO THE HERPESVIRUSES GLYCOPROTEIN B FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; D00818; BAA0699.1; --
 DR PIR; A38478; VGBEL.
 DR InterPro; IPR000234; Glycoprot B.
 DR Pfam; PF00606; Glycoprotein B; 1.
 DR ProDom; PD000693; Glycoprot B; 1.
 KW Signal; Glycoprotein; Transmembrane.
 FT SIGNAL 1 26 POTENTIAL.
 FT CHAIN 27 883 GLYCOPROTEIN B.
 FT DOMAIN 27 724 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 725 744 POTENTIAL.
 FT TRANSMEM 751 771 POTENTIAL.
 FT DOMAIN 772 865 CYTOPLASMIC (POTENTIAL).
 FT CARBOHYD 102 102 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 121 121 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 211 211 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 262 262 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 360 360 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 579 579 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 635 635 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 649 649 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT SEQUENCE 883 AA; 100160 MW; D457DF5178840A7D CRC64;
 Query Match 18.3%; Score 830; DB 1; Length 883;
 Best Local Similarity 27.8%; Pred. No. 2.2e-43;
 Matches 221; Conservative 159; Mismatches 308; Indels 108; Gaps 23;
 QY 57 SHGERGNENRD--SEEQNKNY-----GSPSTPPYR----VCSAGSGVDVFRFQTDHV 104
 DB 34 SHGIAGIDPRDASMDVGKISFSAIGSGAPKEQIQRNIRIPACSPGASVARLAQPRH 93
 QY 105 CPDASDMVH-SGILLIYKQNIIPMERVKRKVVTSTVYNGIYSIDSTNQHTFYKSI 163
 DB 94 CHRHADSTNMTGIAVWVFQNIAPVFNVTLYKHITTTWALFSPQITNEYVTRVPI 153
 QY 164 EPMETEKMDTIYQCNLSRLNTGGNLLTYVDKDDNMVTFLOP-----VDGVTPDV 214
 DB 154 DYHEIVIRDSRGECSKATYHKNFKNFMFFEAYDNDEAEKKLPLVPFSLRSTVSKAFHTTFT 213
 QY 215 KRYGQPELYLPGFWGWSYRRRTTNCCLMDMFARSNPPFFFTATGDTVEMSPFWSG 274
 DB 214 KRHTL-----GYRTSTSDCVVEYLQARSVPDYFGWAGDVTVEISFFYT- 260
 QY 275 EDDHKNKHEKPFWSVNNYK---VVDYQNR-----GTVPGLKTRIFLDREYTLSEKX 327
 DB 261 -----KNTGPRRHSVYRDYRFLFIANYQVRDETGTQIRPPKKNFLTDEQFTIGWDAM 314

QY 328 LKNMSYCPCLTLWKAFYNGIQTEHSGSYHFVANDITASFTTKEDMKERTT-----YHCL- 382
 DB 315 EEKESVCTLSKWEIEVPEAVRVSYKNSYHFSKDMTMTFSSGKQ---PNISRLHLAECVP 371
 QY 383 ---NEEIKAEIEKYYAKVNSTHSHYKDGUKYFKTGGGLYLVMOPLIQNRLD---AKNKLN 436
 DB 372 TIASEAIDGIFARKY---SSTHVRSGDIEYVLGSGGFLIAFQKLMHGLAEWYLEEAQK 428
 QY 437 NETYSRRSRROAESTTDPMMEMTGNGAGEYSSSENSITVAQVOYAYDNLIRINNILEDL 496
 DB 429 NHPRGRERROAQRRTASLQ--SGPOGDRITTHSSATFAMLOFAYDKIOAHVNELIGNL 486
 QY 497 SKAWCREHRAALVMNELSKINPTSMVSMIYNRPVSAKRIGDVISVNCIVVDQTSVSLH 556
 DB 487 LEANCLONRQLIYVHEMKLNPNLSLPGQPSARLLGDIVAVSKCIEIPENIRMQ 546
 QY 557 KSLRLLSASDEKCFSRPPVTFKFM-----NDSTIYKQGLGVNNEILLTTTTL 603
 DB 547 DSMR-MPGDPTWCYTRPVLIFRYSPPSQFSANSTENHNLDILGOLGEHNEILQGRNLI 605
 QY 604 ETCQENTYYFOAKTDMYIKNYEHLKTVPLSSITTLDTFFIALNFTLENVDFKVIELT 663
 DB 606 EPCMINHRRYFLGELYLLYEDYTFVRQVNASEIEVSIFINLNATILEDLDVFPEVYT 665
 QY 664 RDEKRLSNVFDIETMFRYNNYQAR-----VSGLRKDLDLSTNRNQFVDAGSLMDD 716
 DB 666 REEURDTGTLNYDDVVRQNIYKFRDIDTVIRGDRGDAI-----FRAIDAFGNTLGE 720
 QY 717 LG-AVGQTVNNAVSGVATLFFSIVTGFINFKNPFGM-----LMIIVVIGLVFAIYFLT 770
 DB 721 VGKALGTVMVMTAAAAVI---STVSGIASFLSNPFAALGIGIIVVVSIIILGLAFKYVMN 776
 QY 771 KTKIYETAPIKMIYP 786
 DB 777 LKSN-----PVQVLPF 787
 RESULT 27
 VGLB_ILTV6 STANDARD; PRT; 873 AA.
 AC Q02409;
 DT 01-JUL-1993 (Rel. 26, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Glycoprotein B precursor.
 GN GB.
 OS Infectious laryngotracheitis virus (strain 632) (ILTV).
 OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
 OC Alphaherpesvirinae; Infectious laryngotracheitis-like viruses.
 OX NCBI_TaxID=31521;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=92180317; PubMed=1665614;
 RA Poulsten D.J., Adams Burton C.R., O'Brian J.J., Rabin S.J.,
 RA Keeler C.L. Jr.;
 RT "Identification of the infectious laryngotracheitis virus glycoprotein gB gene by the polymerase chain reaction.";
 RL Virus Genes 5:335-347 (1991).
 RN [2]
 RP REVISIONS.
 RA Keeler C.L. Jr.;
 RL Submitted (FEB-1994) to the EMBL/GenBank/DBJ databases.
 CC -!- SIMILARITY: BELONGS TO THE HERPESVIRUSES GLYCOPROTEIN B FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; X56093; CAA39573.1; --

```

DR PIR; S26690; S26690.
DR InterPro; IPR000234; Glycoprotein B.
DR Pfam; PF00606; Glycoprotein B; 1.
DR ProDom; PD000693; Glycoprot_B; 1.
KW Signal; Glycoprotein; Transmembrane.
FT SIGNAL 1 21 POTENTIAL.
FT CHAIN 22 873 GLYCOPROTEIN B.
FT DOMAIN 22 714 EXTRACELLULAR (BY SIMILARITY).
FT TRANSMEM 715 734 BY SIMILARITY.
FT TRANSMEM 741 761 BY SIMILARITY.
FT DOMAIN 762 855 CYTOPLASMIC (BY SIMILARITY).
FT CARBOHYD 92 92 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 111 111 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 201 201 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 252 252 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 350 350 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 569 569 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 625 625 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 639 639 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 873 AA; 99039 MW; E73E7F6D7A0E3AAD CRC64;

Query Match 18.2%; Score 829; DB 1; Length 873;
Best Local Similarity 27.8%; Pred. No. 2.5e-43;
Matches 221; Conservative 158; Mismatches 309; Indels 108; Gaps 23;

Qy 57 SHGERGNENRD--SEQNKNIV-----GSPSTFPYR---VCSASGVGVDFRFQDHY 104
Db 24 SHGIAGIDPRTASMDVGKISFEATGSGAPKEQPIRNIFACSSPTGASVARLAQPRH 83
Qy 105 CPDASDMVH-SEGILLIYKQNIIPFMFRVRYKRVKVTSTVYNGIYSDSITNQTFFYKSI 163
Db 84 CHRHADSTNMTGIAVVFVKQNIAPVFNVTLYKHITVTWALFSRQITNEVYVTRVPI 143
Qy 164 EPWTEKMDTITYQCFNSLRNLGTGNTLYVDRDDINMTVFLOP-----VDGVTPDV 214
Db 144 DYHEIVRIDRSGECSKATYHKNFMPFEAYDNDAEKKLPLVPSLLRSTVSKAFHTTNT 203
Qy 215 KRYGQPELYLEPCGFWGYSYRRRTTVCCLMDMFARSNPPRDFVTATGDTVEMSPFWSG 274
Db 204 KRHOTL-----GYRTSTVDCVBYLQARSVPYDFGMATGDTVEISPPYT- 250
Qy 275 EDDHENKMKHPFVSVINNYK---VVDYQNR-----GTVPGLKTRIFLDREYTLSEKH 327
Db 251 -----KNTTGPRRHSVVDRYFLEIANYQVRDLETGQIRPKPKRNFELTDEQFTIGDAM 304
Qy 328 LKWSYCELTWLKAFYNGIQTEHSGSHFVANDITASFTSKDEMKEFNT-----YHCL- 382
Db 305 EEKESVCTLSKWIEVPEAVRVSYKNSYHFLSKDMTMTFFSGKQ---PFNISRLHLAECVP 361
Qy 383 ---NEEIKAEIKKAKVKNSTHYSKYGDLYKPKTDGGLYLVNWOPLIQNRLLD---AKNKLN 436
Db 362 TIATEADIGTFARKY---SSTHVRSGDIEYVLGGGFLIAFQKUMSHGLAEWYLEEAQRQ 418
Qy 437 NETYSRRSRROAESTTDPMMEMTNGAGGEYSSENSITVAQVQVAYONLRIRINILEDL 496
Db 419 NHPGRGRERROAAGRRTASLQ--SGPQGDRTITHTSSATFAMLOPAYDKIOAHVNELIGNL 476
Qy 497 SKAWCRQHRALVWNLSEKINPTSVMSMYNRPVSAKRIGDVISVNCIIVDQTSVSLH 556
Db 477 LEAWCELQNROLIYVHEMKLNPNSLMTSLFGQVPSARLLGDIIVAVSKCIEPIENIRMQ 536
Qy 557 KSLRLLSASDEKCFSPRPVTFKFM-----NDSTYKQGLGVNNEILLTITL 603
Db 537 DSMR-MFGDPTMCTYRPLVIFRYSSESPEFSANSTENHNDLILGQGEHNEILQGRNLI 595
Qy 604 ETCQENTYFQAOKTMYIYKNYEHKLTVPVLSSTITLDTFIALNFTLLENVDFKVIPLYT 663
Db 596 EPCMINHRRYFLLGENVLLYEDYTFVRQVNASIEEVSIFINLNATILEDLDLDFVPVEVYT 655
Qy 664 RDEKRLNVDFIETMFREYNYAOR-----VSLGRKDLDDUSTNRNQFVDFAGSLMDD 716
Db 656 REELRDTGTLNDDVRYQNIYKRRFRDIDTVIRGDRGDAI-----FRAIADFPNGTLGE 710

```

```

Qy 717 LG-AVGQTVVNAVSGVATLFSSIVTGFINFKPFQGM-----LMIIVVIGLVFAIYELT 770
Db 711 VGRALGTVMVMTAAAVI-----STVSGIASFLSPFAALGIGIAVWVSIIILGLLAFKVMN 766
Qy 771 KTKKIYETAPIKMIYP 786
Db 767 LKSN-----PVQVLPF 777

RESULT 28
VGLB HSV1 STANDARD; PRT; 980 AA.
AC P18050;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 18-OCT-2001 (Rel. 40, Last annotation update)
DE Glycoprotein B precursor (Glycoprotein 14).
GN GB OR GP14 OR 33.
OS Equine herpesvirus type 1 (isolate HVS25A) (EHV-1).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Varicellovirus.
OX NCBI_TaxID=10327;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89279217; PubMed=2543744;
RA Whalley J.M., Robertson G.R., Scott N.A., Hudson G.C., Bell C.W.,
RA Woodworth L.M.;
RT Identification and nucleotide sequence of a gene in equine
RT herpesvirus 1 analogous to the herpes simplex virus gene encoding the
RT major envelope glycoprotein GB."
RL J. Gen. Virol. 70:383-394(1989).
CC -/- SIMILARITY: BELONGS TO THE HERPESVIRUSES GLYCOPROTEIN B FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; D00401; BAA00304.1; ALT_SEQ.
DR PIR; A31241; VGBE2H.
DR InterPro; IPR000234; Glycoprot B.
DR Pfam; PF00606; Glycoprotein B; 1.
DR ProDom; PD000693; Glycoprot_B; 1.
KW Glycoprotein; Transmembrane; Signal.
FT SIGNAL 1 86 POTENTIAL.
FT CHAIN 87 980 GLYCOPROTEIN B.
FT DOMAIN 87 852 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 853 870 POTENTIAL.
FT DOMAIN 871 980 CYTOPLASMIC (POTENTIAL).
FT CARBOHYD 165 165 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 275 275 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 380 380 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 423 423 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 497 497 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 514 514 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 515 515 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 560 560 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 727 727 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 749 749 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 980 AA; 109931 MW; 94B69AF882FB6AC2 CRC64;

Query Match 18.1%; Score 824; DB 1; Length 980;
Best Local Similarity 26.9%; Pred. No. 5.9e-43;
Matches 252; Conservative 155; Mismatches 379; Indels 150; Gaps 31;

Qy 32 VTSPPNATWTSPLTGHGTHDSHGERGNENRDEBQNKNIYCSPTFPVRYVC-SA 90
Db 83 VRAVPTSPGPTSTPTS--MSTH--SHG-----TVDPTLLPTDPDLRLAURE 127
Qy 91 SGV-----GD-----VFRFQDHYVCPDAS-DMVHSEGILLIYKQNIIPFMFRVK 134

```

```

Db 128 SGILAEADGDFYTCPPPTGTVVRIEPPRTCPKFDLGRNFTGEGIAVIFKENIAPYKFRANV 187
Qy 135 YRKVVTTSTVYNGIYSDSIHQHTFYKSIPEWETEK---MDTIYQCENSLRLNTGGNLL 190
Db 188 YKDIWTVRWKGYSHTSLSDR---YNDRVVSVSEIFGLIDSKGKSKAB-----236
Qy 191 TYVDRDDINMTVFLQVGVDPDKVRYGSGQPELYLE-----PGWF-WGSYR 235
Db 237 -YL-RDNIMHAYHDEDEVELDLR---PSQLRGARAWQTTNDTTSYVGWMPRHY- 289
Qy 236 RRTVNCCLMDMFAKSNPPDFVTATGDTVMSPFW-----SGEDDHENKMKHEKPMFVSV 291
Db 290 TSTSYNCIVEEARSVPYDSFALSTGDIIVYASPFYGLRAAARIHNSYAQER---FRQ 346
Qy 292 INNYKVVDPYQNGRTVPLGKTRIFLDREEVTLSEWKLKNMSYCPPLTKAFNGIQTGHS 351
Db 347 VEGYPRDLDSKLQAEPPVTKNFITTPHVTVSNWNTKKEVEACTLTKWKEVDELVRDEFR 406
Qy 352 GSYHFVANDITASFTTSKEDMK-BENTTYHCLNEEIKAEIEKKYAK-VNSTHSGYDGLKY 409
Db 407 GSYRTFIRSISSFTSNTTQFKLESAPLTCVSKAEKAIDSIIYKKQYESTHVFSGDVEY 466
Qy 410 FKTDGGLYLWQPLI-----QNRLLDAKNKL-----NNETYSRRSRQAEST 451
Db 467 YLARGGFLIAFRPMLSNELARLYNELVRSNRTYDLKNLLPNANNNNTTTRRSLSLV 526
Qy 452 TDPMMEMTG-----NGAGEVSS-----NITVAQVQYAYDN 484
Db 527 PEQPTQGVHREQLHLRKRABATAGTSSNTAKOLELITTSIEFAMQLQFAYDH 586
Qy 485 LRIRINILEDLSKAWCEQRAALVWNELSKINPTSVMSIYNNRPVSAKIGDIVSYSN 544
Db 587 IQSHVNEMLSRITATWCTLQNKERTLWNEWKINPSAIVSATLDERVAARVLGDVIAITH 646
Qy 545 CIVDQTSVSHKSLRLLSASDEKCFRPPVTF---KFMNDSTIYKGLGVNNEILLTTT 601
Db 647 CAKI-EGNVYLQNSMR--SMDSNTCYRPPVTFITTKNANNRSGIEGQGEENEIFTERK 703
Qy 602 YLETCQENTYVFOAKTMYIKYNEHLKTVPLSSITTLDTFIALNFLENVDFKVL 661
Db 704 LIPECALNOKRYFKGKYVYENYTFVRKVPPTIEIVSYVNLNLTLEDREFLPLEV 763
Qy 662 YTRDEKRLSNVDFIETMREYNVYVQSVGLRKLDDLSLRNQVQVDAFGSLMDDLGAVG 721
Db 764 YTRAELEDGLDYSEIORNQLHALREFYDI--DSVNVNDTAVIMQGIASFFKGLGKVG 821
Qy 722 QTVNAVSGVATLFSSIVTGFINFKNPFGLMLIIIVIGLFAIYFLTKTKIYETAPI 781
Db 822 EAVGLVLGAAGAVVSTVSGIASFLNPFGLAIGLLVIAGLVAFAFFAYRYVMQIRSNPM 881
Qy 782 KMIYPEIDKLKEREKSEIAPISE-----EELERIVLAMIHQHSHMETKTRK 830
Db 882 KALPYITTKALKNKAITSYGNEEDGDFDEAKLEAEAREMIKYMSVMVSALEKQEKKAIK 941
Qy 831 DPKDS-----ILTRAQMLRKSG--YSLNKAESVE 860
Db 942 --KNSGVGLIASVSKLARRRGPKYTRLQNDTME 975

```

RESULT 29

```

VGLB_HSVEA
ID_VGLB_HSVEA STANDARD; PRT; 980 AA.
AC P18551;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Glycoprotein B precursor (Glycoprotein 14).
GN GB OR GP14 OR 33
OS Equine herpesvirus type 1 (strain AB1) (EHV-1).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Varicellovirinae.
NCBI_TaxID=10328;

```

```

RN SEQUENCE FROM N.A.
RA Bonass W.A., Elton D.M., Stocks J.M., Killington R.A.,
RA Meredith D.M., Halliburton I.W.,
RL Submitted (JUL-1990) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: BELONGS TO THE HERPESVIRUS GLYCOPROTEIN B FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; M36298; AAA46068.1; --
DR InterPro; IPR000234; Glycoprot B.
DR Pfam; PF00606; Glycoprotein B; 1.
DR ProDom; PD000693; Glycoprot B; 1.
KW Glycoprotein; Transmembrane; Signal.
FT SIGNAL 1 86
FT CHAIN 87 980
FT DOMAIN 87 852
FT TRANSMEM 853 870
FT DOMAIN 871 980
FT CARBOHYD 165 165
FT CARBOHYD 275 275
FT CARBOHYD 380 380
FT CARBOHYD 423 423
FT CARBOHYD 497 497
FT CARBOHYD 514 514
FT CARBOHYD 515 515
FT CARBOHYD 560 560
FT CARBOHYD 727 727
FT CARBOHYD 749 749
SQ SEQUENCE 980 AA; 109736 MW; 264D273CED70ESA1 CRC64;

```

Query Match 18.1%; Score 824; DB 1; Length 980;

Best Local Similarity 26.7%; Pred. No. 5.9e-43;

Matches 248; Conservative 161; Mismatches 385; Indels 134; Gaps 31;

```

Qy 32 VTPSPNTATSTESPLTGHYTHSDSHGCGNNENRDSSEQNKNIYSPSPFPYRVC-SA 90
Db 83 VRAVPTTSPSPSTPTS--MSTH--SHG-----TVDPDLLPTETDPLRLAVRE 127
Qy 91 SGV-----GD-----VFRQTHVCPDAS-DMVHSEGLILLIYKQNIIPWFRVK 134
Db 128 SGILAEADGDFYTCPPPTGTVVRIEPPRTCPKFDLGRNFTGEGIAVIFKENIAPYKFRANV 187
Qy 135 YRKVVTTSTVYNGIYSDSIHQHTFYKSIPEWETEK---MDTIYQCENSLRLNTGG 187
Db 188 YKDIWTVRWKGYSHTSLSDR---YNDRVVSVSEIFGLIDSKGKSKABEYLR----D 240
Qy 188 NLLTYV---DRDDINMTVFLQVGVDPDKVRYGSGQPELYLEPGWF-WGSYRRRTTVNCE 243
Db 241 NIMHAYHDEDEVELD--LVPSKEATPGARAWQTTNDTTSYVGWMPRHY-TSTSVNCI 297
Qy 244 LMDMFAKSNPPDFVTATGDTVMSPFW-----SGEDDHENKMKHEKPMFVSVINNYKVD 299
Db 298 VEEVEARSVPYDSFALSTGDIIVYASPFYGLRAAARIHNSYAQDS---FRQVEGYRPRD 354
Qy 300 YQNGRTVPLGKTRIFLDREEVTLSEWKLKNMSYCPPLTKAFNGIQTGHSYHVPAN 359
Db 355 LDSKLQAEPPVTKNFITTPHVTVSNWNTKKEVEACTLTKWKEVDELVRDEFRGSYRFTIR 414
Qy 360 DITASFTTSKEDMK-BENTTYHCLNEEIKAEIEKKYAK-VNSTHSGYDGLKYFKTDGGLY 417
Db 415 SISSTFISNTTQFKLESAPLTCVSKAEKAIDSIIYKKQYESTHVFSGDVEYILARGGFL 474
Qy 418 LVMQPLI-----QNRLLDAKNKL-----NNETYSRRSRQAESTTDPMMEMT 459
Db 475 IAFRPMLSNELARLYNELVRSNRTYDLKNLLPNANNNNTTTRRSLSLVPEPQPTQD 534

```

```
QY 460 G-----NGAGGEYSSE-----NSITVAQYQAYDNLIRINNI 492
Db |-----|-----|-----|-----|-----|-----|
535 GVHREQLHLRHKRAVEATAGTSSNVTAQLELILKTTSSIEFAMQFADHIOQSHVEM 594
QY 493 LEDLSKAWCREQRAALVWNELSKINPTSVMSIYNRPVSAKRIGDVISVNCIVVDQTS 552
Db |-----|-----|-----|-----|-----|-----|
595 LSRIATAWCTLQNKERTLMNEVKINPSAIVSATLDERVAARVLGDVIAITHCAKI-EGN 653
QY 553 VSLHKSRLLSASDEKCFSPBPVTF---KEMNDSTIYKGGOLGVNNEILLTTTILETCQEN 609
Db |-----|-----|-----|-----|-----|-----|
654 VYLQNSMR---SMDNTCYSPBPVTFITTKNANRGSIEGOLGEBENEIFTERKUIEPCALN 711
QY 610 TETVFOAKTDMYIKNYEHLKTVPLSSITLDTFIALNFTLLENVDPKVIETLREKRL 669
Db |-----|-----|-----|-----|-----|-----|
712 QKRYFKGKGVYVYENFTVRKVPPTETIEVISYVELNLILLDEPLPLEVTRAELD 771
QY 670 SNVPDIETMPREYNYYAQRVSGRLKOLLDLSTNRNQPVDAGSLMODLGAVGTVNVAS 729
Db |-----|-----|-----|-----|-----|-----|
772 TGLLDYSEIQRRNQHALREFYDI--DSVNVNDTAVIMQGIASFFKGLGKVGGEAVGLVL 829
QY 730 GVATLFSSIVTGFNFINKPFGGMLIIVIGVLFAIYFLTKTIYETAPIKMIYPEID 789
Db |-----|-----|-----|-----|-----|-----|
830 AAGAVVSTVSGIASFLNPFGLAIGLVIAAGLVAAFFAYRYVMQIRSNPMKALYPITT 889
QY 790 KLKERECKSEIAPISE-----EELERIVLAMHIHQONSHMETKTRDKPKDS--- 835
Db |-----|-----|-----|-----|-----|-----|
890 KALKNKAKTSGYQNEEDGSDDFDEAKLEAREMIKYSMWSALEKQEKKAIK--KNSGVG 947
QY 836 -LITRAQNMLRKRSQ--YSNLKNAESVE 860
Db |-----|-----|-----|-----|-----|-----|
948 LIASNVSKALRRRGPKYTRLQNDTWE 975

RESULT 30
VGLB HSA1 STANDARD; PRT; 933 AA.
AC Q0463;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Glycoprotein B precursor.
GN Ul27.
OS Herpesvirus ateles (type 1 / strain Lennette).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae.
OX NCBI_TaxID=35243;
RN [1]
SEQUENCE FROM N.A.
RX MEDLINE=93228440; PubMed=8385913;
RA Eberle R., Black D.;
RT "Sequence analysis of herpes simplex virus gB gene homologs of two
RT platyrrhine monkey alpha-herpesviruses.";
RL Arch. Virol. 129:167-182(1993).
CC -|- SUBUNIT: DIMER, PROBABLY LINKED BY DISULFIDE BONDS.
CC -|- SUBCELLULAR LOCATION: Integral membrane protein.
CC -|- MITOCHONDRIAL: GB IS THE ONLY GLYCOPROTEIN THAT IS KNOWN TO BE
CC REQUIRED FOR VIRAL GROWTH.
CC -|- SIMILARITY: BELONGS TO THE HERPESVIRUSES GLYCOPROTEIN B FAMILY.
CC -|- SIMILARITY: IT IS UNCERTAIN WHETHER MET-1 OR MET-30 IS THE INITIATOR.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M95785; AAA43839.1; --
DR InterPro: IPR000234; Glycoprot_B.
DR Pfam: PF00606; Glycoprotein_B; 1.
DR ProDom: PD000693; Glycoprot_B; 1.
KW Glycoprotein; Transmembrane; Signal.
```

```
FT SIGNAL 1 31 POTENTIAL.
FT CHAIN 32 933 GLYCOPROTEIN B.
FT DOMAIN 32 733 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 754 769 POTENTIAL.
FT TRANSMEM 775 795 POTENTIAL.
FT TRANSMEM 798 818 POTENTIAL.
FT DOMAIN 819 933 CYTOPLASMIC (POTENTIAL).
FT CARBOHYD 107 107 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 161 161 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 418 418 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 450 450 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 697 697 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 747 747 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 933 AA; 101492 MW; 13A6CF88B5A18DF CRC64;

Query Match 18.1%; Score 821; DB 1; Length 933;
Best Local Similarity 26.4%; Pred. No. 8.5e-43;
Matches 224; Conservative 164; Mismatches 378; Indels 84; Gaps 20;

QY 30 TGVTSPPNTATWSTESPLTGHYGDTHDSSHGERGNN---ENRDSBEONKNIYSGSPSTFPY 85
Db |-----|-----|-----|-----|-----|-----|
80 SGPEP-----GPDPRPGPGGSGRRRSGPCNGTRSAARQLRESLRRIQAEYASAF 133
QY 86 RYCSASGVGDVFRFQTDHVCPPD-ASDMVHSEGLLIYKQNIIPFMFRVKRYKVVITSTV 144
Db |-----|-----|-----|-----|-----|-----|
134 YVCPPTGATVQFEBEPCPDVAAGKNFTAGIAVVFKENIAPYKFTATKYKEITVSQT 193
QY 145 Y-----NGIYSOSITNQHTFYKSIPEWE--TEKMDTIYOCFNSLRNLTGNNLTYYVD 194
Db |-----|-----|-----|-----|-----|-----|
194 WQGSRYLQLTGLYND-----RAPVPFSEITDLINGKGRCSRDVTVTRSORRTAYD 244
QY 195 RUDINMTVFLOPVDGVTDPVKYSGOPELYLBPQFWSYRRTTVCNCELMDFMFAFNSNP 254
Db |-----|-----|-----|-----|-----|-----|
245 GDEWGREVALVPAKTSPTNSRGWYTTDRVY-APNAHAGFYKGTGVNCEIVEEMEARSAFP 303
QY 255 FDFPVTATGDTVEMSPFSGEDD--HENKMHKEKPFVSVINNY-KVVDYQNRCITVPLGKT 311
Db |-----|-----|-----|-----|-----|-----|
304 YDSFVLATGEFVYVASFSGFSEDARERNRYAPDRFQVDGFPYPRDLDSGQRAATPV--V 361
QY 312 RIFLDREEVTLSEKHLKNMSYCPULWKAFYNGIOTHSYGVHVFVANDITASFTTSKED 371
Db |-----|-----|-----|-----|-----|-----|
362 RNLLATPTFTVGMDWKPRPNVCSVTKWQVBEEMVRAEYGSAPRFTSAALSATFTS--- 417
QY 372 MKEFNNTYH-----CLNEEIKAEIEKKA-KVNSTHYSKYGDLKYFKTDCGLVLVMO 421
Db |-----|-----|-----|-----|-----|-----|
418 ----NLTQPPPELIEHSDCVAREAAESIETAIYARRYNASHVRVGGVQVYLAAGFFLAFQ 473
QY 422 PLIQNRLLDAKNLNNETYSRSRROAESTTDPMMEMTNGAG--GEYSSENSITVAQVQ 479
Db |-----|-----|-----|-----|-----|-----|
474 PLLSNSLAEMVRR---EALLGRSGDLAALAPPVAAAPASGAGPRCTISTTQTVFARLQ 530
QY 480 YAYDNLIRIRINILEDLSKAWCREQRAALVWNELSKINPTSVMSIYNRPVSAKRIGDV 539
Db |-----|-----|-----|-----|-----|-----|
531 FTVDHTQKHVNEMLGRIAAAWQQLQBELVWNEARKLNPGATASATATVGTFRVARMGLDV 590
QY 540 ISVNCIVVDQTSVSLHKSRLLSASDEKCFSPBPVTFEKFMDNSTIYKGGOLGVNNEILLT 599
Db |-----|-----|-----|-----|-----|-----|
591 MAVSTCIPSPDNVIMQNSMR-IPGDPKTCYARPLVSFRYTDDEGLVEGOLGSDNEIRLE 649
QY 600 TTVLETQENETTYFOAKTDMYIKNYEHLKTVPLSSITLDTFIALNFTLLENVDPKVI 659
Db |-----|-----|-----|-----|-----|-----|
650 QNNVECTVGHKRYFVFGGVYVFEYAYSHQVSRAADVPVWSTFVDNLNLTMLDEHFLPL 709
QY 660 ELYTRDEKLSNVFDIETMFPREYNYYAQRVSGRLKOLLDLSTNRNQPVDAGSLMODLGA 719
Db |-----|-----|-----|-----|-----|-----|
710 EVVTRREIKSDGLDYAEVQRRNQLHALRFSIDIRIMND-SANA-ALMAGLARFFDQMGD 767
QY 720 VQOTVNVAVSGVATLFPSSIVTGFNFIKNPFGGMLIIVIGVLFAIYFLTKTIYETA 779
Db |-----|-----|-----|-----|-----|-----|
768 AGAIGRAVLGVTEGLISVVGSSFLSNPFGALAVGLLVLAAGLVAAFFAMRIMRLAN 827
QY 780 PIKMIYPEI-----DKLKERECKSEIAPISEELE-----RIVLAMHI 817
Db |-----|-----|-----|-----|-----|-----|
```


KW Glycoprotein; Transmembrane; Signal.
 FT SIGNAL 1 86
 FT CHAIN 87 979
 FT DOMAIN 87 979
 FT TRANSMEM 852 869
 FT DOMAIN 870 979
 FT CARBOHYD 165
 FT CARBOHYD 275
 FT CARBOHYD 380
 FT CARBOHYD 423
 FT CARBOHYD 497
 FT CARBOHYD 514
 FT CARBOHYD 515
 FT CARBOHYD 560
 FT CARBOHYD 727
 FT CARBOHYD 749
 SQ SEQUENCE 979 AA; 110331 MW; 9A19866B791C5B36.CRC64;

Query Match 17.1%; Score 777.5; DB 1; Length 979;
 Best Local Similarity 26.3%; Pred. No. 4.2e-40;
 Matches 247; Conservative 159; Mismatches 377; Indels 155; Gaps 34;

QY 32 VTSPPNTATWSTESPLTGHYTHDSHGCGNNEIRDSEEQKNLYGSPSPFPYRVC-SA 90
 Db 83 VRAPVTPSPPTSTPTS--MSTH--SHG-----TVDPTLLPTETDPLRLAVRE 127
 QY 91 SGV-----GD-----VRFOTDHCVPDAS--DMVHSEGLIILYKONIIIPFMRVRK 134
 Db 128 SGILAEEDGDFYTCPPPTGTVVRISPPRTCPKFDLGRNFTGIAVIFKENTAPYFRANV 187
 QY 135 YRKVTTSTVYNGIYSDSITNHTYKSIPEWTEK-----MDTIYQCENS---LRLNTGG 187
 Db 188 YKDIWTVTRVMKGYSHTSLSDR--YNDRVVPVSEIEFGLIDSKGCKSKAEYLA---D 240
 QY 188 NLLTV---DRDDINMTVLPQVGDVTDVKRYGQPELYLEPGWF--NGSVRRRTTVNCE 243
 Db 241 NIMHAYHDDDEVELD--LVPSKFPATPGARAWQTNDTTSYVGMWPRHY--TSTSVNCI 297
 QY 244 LMDFEARNPPDFVTATGDTVEMSPW----SGEDDHENKMKHKPWFVSVINNYKVD 299
 Db 298 VEEVARSVYFDSFALSTGDIYVSPYGLRAAARIENHNSVAQER---FRQVEGYRPRD 354
 QY 300 YONRTVPLGKTRIFLDREETYLSWEKHLKNMSYCPJTLWKAFYNGIQTHSGSYHFVAN 359
 Db 355 LDSKLQAEPEVTKNFITPHVTVSNWTEKKVEACTLTWKVEVDLVRDFRGSYRFTIR 414
 QY 360 DITASFTTSKEDMK--EFNTHYHCLNEEIKAEIKKYAK--VNSTHSGYGLDYKFTDGGLY 417
 Db 415 SISSYFISNTTQFKLESAPLTCVSKAEKAIDSYKKQYESTHYVFGDVEYLYLARGGFL 474
 QY 418 LWQPLI-----QNRLLDAKNKL-----NNETYSRRSRROAESTTDPMEWT 459
 Db 475 IAFRMLSNELARLYNLVLSNRNYDLYKLLNPNANNNNNTTTRRRSLLSVPEQPQTD 534
 QY 460 G-----NGAGGEYSSE-----NSITVAQVQVAYDNLRIINNI 492
 Db 535 GVHREQLHLRLHRAVEATAGTDSNVTAQKLELTKTSSIEFAMQLQFAYDHQSHVNM 594
 QY 493 LEDLSKANCREQHRAALVWNLSEKINPTSVMSMYNRRPVSAKRIGDVISVNCIVDQTS 552
 Db 595 LSRIATAWCPQLQNKERPLNEMVKITPSAIVSATLDERVAARVLGDVIAITHCAKI--EGN 653
 QY 553 VSLKSLRLLSADSEKCFSPDPVTF--KFVNDSTIYKGLGVNNEILLTLYLETQEN 609
 Db 654 VYLQNSMR--SWDSNTCSRPPTVTITKNANNRSGIEGQGEENEIFETERKLEPCALN 711
 QY 610 TEYFQAKTDYIYKNYELHTKVPSSITTTDLTFIALNFTLLENVDKVIETYRDEKRL 669
 Db 712 QKRYFKGCKEYVYENYFVRKVPTEIEVISTYVELNLTLLEDEFLPLEVYTFEALED 771
 QY 670 SNVFIETMFREYNYAQRVGLRKDLDDLSTNRNQFVD-----AFGLMDDLGA 719
 Db 772 TGLDYSBIORRNQHLALR-----FYDIDSVVN--VDNTAVIIRGSPAFSAWVKWR 822

QY 720 VQGVNAVSGVATLFSFIVTGFINFKNPFGLMIIWIGVLFAIYFLTKTKIYETA 779
 Db 823 PWERSFSARGAVV-----STVSGIACFLNNPFGGLIAGLVAAFAFYVYVQIRSN 878
 QY 780 PKIMYPIEIDKLKEREGKSEIAPISE-----EELERIVLAMHIHOONSHMETKT 828
 Db 879 PMKALYPITTKALKNKAKTYSQNEEDDGSDFDEAKLEAREMIKYMVMVSALEKQEKKA 938
 QY 829 RKDPKDS-----ILFPAQNMLKRSG--YSNLKNAESVE 860
 Db 939 IK--KNSGVGLIASNVSKLALRRRGPKYTRLQONDIME 974

RESULT 34
 VGLB PRIVIF
 ID VGLB PRIVIF STANDARD; PRT; 913 AA.
 AC P08355;
 DT 01-AUG-1988 (Rel. 08, Created)
 DT 01-AUG-1988 (Rel. 08, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Glycoprotein GII precursor.
 OS Pseudorabies virus (strain Indiana-Funkhauser / Becker) (PRV).
 OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
 OC Alphaherpesvirinae; Varicelloviruses.
 OX NCBI_TaxID=31523;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=87284141; PubMed=3039163;
 RA Robbins A.K., Dorney D.J., Wathen M.W., Whealy M.E., Gold C.,
 RA Watson R.J., Holland L.E., Weed S.D., Levine M., Glorioso J.C.,
 RA Enquist L.W.;
 RT "The pseudorabies virus gII gene is closely related to the gB
 glycoprotein gene of herpes simplex virus.";
 RL J. Virol. 61:2691-2701 (1987).
 RN [2]
 RP SEQUENCE OF 847-913 FROM N.A.
 RX MEDLINE=89279298; PubMed=2543777;
 RA Simon A., Mettenleiter T.C., Rziha H.J.;
 RT "Pseudorabies virus displays variable numbers of a repeat unit
 adjacent to the 3' end of the glycoprotein gII gene.";
 RL J. Gen. Virol. 70:1239-1246 (1989).
 RN [3]
 RP EXPORT PATHWAY.
 RX MEDLINE=90219190; PubMed=2157862;
 RA Whealy M.E., Robbins A.K., Enquist L.W.;
 RT "The export pathway of the pseudorabies virus gB homolog gII involves
 oligomer formation in the endoplasmic reticulum and protease
 processing in the Golgi apparatus.";
 RL J. Virol. 64:1946-1955 (1990).
 CC - SUBUNIT: DIMER, PROBABLY LINKED BY DISULFIDE BONDS.
 CC - SIMILARITY: BELONGS TO THE HERPESVIRUSES GLYCOPROTEIN B FAMILY.
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (see <http://www.ebi.ac.uk/ebis-sib.ch/>)
 or send an email to license@sib-sib.ch.
 DR EMBL; M17321; AAA47465.1; -;
 DR EMBL; D00464; BAA00359.1; -;
 DR PIR; A29159; VGBEPS.
 DR InterPro; IPR000234; Glycoprot B.
 DR Pfam; PF00606; Glycoprotein B; 1.
 DR ProDom; PD000693; Glycoprot_B; 1.
 Glycoprotein; Transmembrane; Signal.
 SIGNAL 1 40
 FT CHAIN 41 913 GLYCOPROTEIN GII.
 FT DOMAIN 41 750 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 751 819 POTENTIAL.
 FT DOMAIN 820 913 CYTOPLASMIC (POTENTIAL).

```
FT CARBOHYD 151 N-LINKED (GLCNAC. . .) (POTENTIAL)
FT CARBOHYD 261 N-LINKED (GLCNAC. . .) (POTENTIAL)
FT CARBOHYD 441 N-LINKED (GLCNAC. . .) (POTENTIAL)
FT CARBOHYD 516 N-LINKED (GLCNAC. . .) (POTENTIAL)
FT CARBOHYD 573 N-LINKED (GLCNAC. . .) (POTENTIAL)
FT CARBOHYD 633 N-LINKED (GLCNAC. . .) (POTENTIAL)
FT CARBOHYD 697 N-LINKED (GLCNAC. . .) (POTENTIAL)
SQ SEQUENCE 913 AA; 100234 MW; 5D560D235E956437 CRC64;

Query Match 17.1%; Score 775.5; DB 1; Length 913;
Best Local Similarity 27.0%; Pred. No. 5.1e-40;
Matches 235; Conservative 165; Mismatches 388; Indels 81; Gaps 26;

Qy 26 SIATGTVSPNTATWSTESPLTCHYGHDSHGCGNENRDSBEQKNLYGSPSPFPY 85
Db 85 SLEIEAFSP-----GPSEAP-DGEYGLDARTAVRAAATERDR-----F 123

Qy 86 RVCSAGSGVDVFRPQTQDHVCPDASD-MVHSGILLIYKQNIIPFMFRVKYKRVVTTSTV 144
Db 124 YVCPSPSGSTVVRLEPEQACEPYSGQRNFTBGIAVLFKENIAPHKFAHIYKXVITTV 183

Qy 145 YNGIYSOITNQHTFYKSIEMWE-TEKMDTIYQCFNSLRNTGNNLLTYVDRDDINMTVF 203
Db 184 WSGSTYAAITNRFTDRVPVQVEITDIDRRGKCVSKAEYVRNNHKVTAFFDRDENPVEVD 243

Qy 204 LQPDVGVTPDKVYSGOPELVLEPGWFGSVYRRRTVNCCLMDMFABSNPPDFEFTVATG 263
Db 244 LRPRNLALGTRGWHHTNDYTKIG-AAGFYHTGTGVCNICEVEARSVPYDYSFALSTG 302

Qy 264 DTVMSPFWS-GEEDHENKMEKFPWFSVINNYKVQVQNRGTPLGKTRIFIDREBYTL 322
Db 303 DIVTMSFPYGLREGAHEGHIYAGPRFQOVHEHYPIDDLSDRLRASESVTRNFLTPTFTV 362

Qy 323 SWEKHLNMSYCPITLWKAFFNGIQTE-HSGSYHFVANDITASPTT--SKEDMKFENYTY 379
Db 363 AWDWAPKTRRYCSLAKWEAEEMTRDETROGSRFTSRALGASVDSVDTQLDRVHLG- 421

Qy 380 HCLNEEIKAEYK-KAVNSTHSGYKGL-KYKFTDGLVLMQPLIQNLRLDA-KNKLN 436
Db 422 DCVUREASEALDAILYRRYNSHTVLADRPPEVILARGGVFAFRPLISNELAQYARELE 481

Qy 437 N-----ETYSRRSRQ--AESTDPMEMTNGAGGEYSSENSITVAQOVAYD 483
Db 482 RLGLAGVGPAPAAARRARRSPGAGTPEP---PAVNGTGLRITTTGSAEFARLQTYD 538

Qy 484 NLRIRINILDLKACWCRQHRALVNLSEKINPTSVMSMIYVRPVSAKRIGDVTSVS 543
Db 539 HIQAVHNDMLGRIAAACELQNDRTLWSEMSRLNPSAVATAALGQVSARMGLGDVMAIS 598

Qy 544 NCIVVDQTSVSLKSLRLSLASDEKCFSPRPVTPKFMNDSTIYKQGVNNEILLTLYL 603
Db 599 RCVEV-RGGVTVQNSMR-VGERGTCYSRPLVTEH--NGTVIEGGQGDNDNELLISRDLI 655

Qy 604 ETCOENTYFPAKTMVYIKNYEHLKTVPLSSITLTDTFIALNFTLLENVDKVIELYT 663
Db 656 EPCGTGNHRRYFKLGSGYVYEDYNYVRMVEPE--TISTRTVNLTLLEDRFLPLEVYT 713

Qy 664 RDEKRLSNVDIETMFREYNYA-----QRVSGLRKDLDLSTNRNQFVDAGFSLMDDL 717
Db 714 REELADTGLLDYSEIQRRNQLHALKFYDIDRVKVDHNVLLRGIANFF-----QGL 765

Qy 718 GAVGQTVNNAVSGVATLFFSSITGTFINFIKNPFGMLMIIVIGVLPAIFYLTKTKIYE 777
Db 766 GDVGAAGVKVLGATGAVIYAVGGWVSFLNPSFGALGILVLVLGVLAAFLAVRHISRLR 825

Qy 778 TAPIKMTYPIDLKEREGKSEIAPISEEBLER-----IVLAMHIHQONSHMETKT 828
Db 826 RNPWKALYPVTTKTKEDGVDE-GDVDEAKLDQARDMIRYMSIVSAL---EQEHRKARK 881

Qy 829 RKPDKSILTRAQNMRLKRSGYSLNKNAE 857
Db 882 NSGPA-LLASRVGAMATRRRRHYQRLESD 909
```

```
RESULT 35
VGLB_HSVBP STANDARD; PRT; 928 AA.
ID VGLB_HSVBP
AC P17471;
DT 01-AUG-1990 (Rel. 15, Created)
DT 01-AUG-1990 (Rel. 15, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Glycoprotein I precursor (Glycoprotein GVP-6) (Glycoprotein 11A)
DE (Glycoprotein 16) (Glycoprotein G130).
GN GI.
OS Bovine herpesvirus type 1 (strain P8-2).
OC Viruses; dsDNA viruses, no RNA strain; Herpesviridae;
OC Alphaherpesvirinae; Varicellovirus.
OX NCBI_TaxID=10324;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89020821; PubMed=2845660;
RA Misra V., Nelson R., Smith M.;
RT "sequence of a bovine herpesvirus type-1 glycoprotein gene that is
homologous to the herpes simplex gene for the glycoprotein gB.";
RL Virology 166:542-549(1988).
CC -!- SUBUNIT: DIMER, PROBABLY LINKED BY DISULFIDE BONDS.
CC -!- SIMILARITY: BELONGS TO THE HERPESVIRUSES GLYCOPROTEIN B FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M23257; AAA46013.1; -.
DR PIR; A31166; VGBEBG.
DR InterPro; IPR000234; Glycoprot B.
DR Pfam; PF00606; Glycoprotein B; 1.
DR ProDom; PD000693; Glycoprot_B; 1.
KW Glycoprotein; transmembrane; Signal.
FT SIGNAL 1 67
FT CHAIN 68 928 GLYCOPROTEIN I.
FT DOMAIN 68 780 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 781 801 POTENTIAL.
FT TRANSMEM 804 824 POTENTIAL.
FT DOMAIN 825 928 CYTOPLASMIC (POTENTIAL).
FT CARBOHYD 105 105 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 153 153 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 442 442 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 484 484 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 579 579 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 637 637 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 703 703 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CONFLICT 409 409 S -> T (IN REF. 1; AAA46013).
FT CONFLICT 673 673 G -> P (IN REF. 1; AAA46013).
SQ SEQUENCE 928 AA; 102177 MW; B47982224FCD769D CRC64;

Query Match 15.9%; Score 722; DB 1; Length 928;
Best Local Similarity 26.2%; Pred. No. 1e-36;
Matches 240; Conservative 156; Mismatches 427; Indels 94; Gaps 26;

Qy 2 AGSLKRGSVLALWLYQVALYSLSIAETCVT-----SPNTATWSTESPLTGHY 51
Db 41 AGGARAALAAALLWATWALLLAAAPAGRAPATTPAPPPEEAASPAFPSPGPGDDA 100

Qy 52 GTHDSSHGGERGNENRDSBEQKNLYGSPSPFPYRVCSAGSGVDVFRFQTDHVCDD 110
Db 101 ASPDNSTDVRAALRLQAAGENSRRF-----VCPSPSGATVVR LAPARCPCEVGLG 151

Qy 111 MVHSEGIILYKQNIIPFMFRVKYKRVVTTSTVYNGIYSDSITNOHTFYKSIEMWE-TE 169
Db 152 RNYTEGIGVYKENIAPYTFKAIYYKXVITTVTTWAGSTYAAITNQYTRVPVGMGEITD 211

Qy 170 KMDTIYQCFNSLRNTGNNLLTYVDRDDINMTVFLQPDVGVTPDKVYKYSQPELYLEPGW 229
```

```

Db 212 LVDDKWRCLSKAABYLRSGKVVAFDRDDDPWEPAPLKPARLSAPGVGRWHHTDDVYALG- 270
Qy 230 FWGYSRRRTTVNCELMDMFARGNPPDFPVATGDTIVEMSPWS--GEDDHENKMEKWF 288
Db 271 SAGLYRTGTSVNCIVEEVARSVYDPSFALSTGBIIYMSPPYIGREGAHEHTSYSPER 330
Qy 289 VSVINNYKVVDYQNRGTVPGLKTRIFLDRREYTLISWEKHLKMSYCPILTMKAFYNGIQT 348
Db 331 FOQIEGYKRDMMATGRRLKEPVSRRNFLRQHVTVAWDVPKRVKNCVSLAKWREADMLRD 390
Qy 349 EHSYSYHVFANDITASFITTSKEDMKEFNNTY--HCLNEELKABIEKYY--AKVNSTHSGVD 406
Db 391 ESRGNFRFRFARLSATFVSDSHTFALQNVPLSDCVIEEAAEAVERVRYRNGTHTVLGSG 450
Qy 407 LKYFKTDGGLYLWVOPLOINRLLDAKNKL-----NNETYS-----RRRRQAE 449
Db 451 LETYLARGGFVAFRPMLSNEL--AKLYLOELARNGTLEGLFAAAAPKPGRRARRRR 508
Qy 450 STTDPMMETNG-----AGBYSSENSITVAQVAYDNLRIINNILEDLSKAWC 501
Db 509 LRPAP-----GRQARRRRHAGGRVTVSLAEFAALQFTHDTRTSEHHV--HRLASPMW 562
Qy 502 RQHRRAALVWNLKINPTSMYINRPVSARKIG-----DVI--SVNCIVVDQTSVSL 555
Db 563 LLQNKERALLWAEAKLNPASAAASALDRPPRACWGTGHRDVLPRAGRGALFIENSARA 622
Qy 556 HKSLRLLSASDEKCFSRPPVTFKFNWDSTIYKGLGVNNEILLTTTLYLETCOENTYYFQ 615
Db 623 RR--RL-----QPRFLSFGNESEPVQGEDNELLPGRELVEPCPTANKHRYFR 671
Qy 616 AKTDMYIYKNYBHLKTVPLSSITTTDLTFIALNFTLLENVDFKVIETYTRDEKRLSNVFDI 675
Db 672 FGADVVYVYENYAVRRVPLAELEVISFTVDNLNTVLEDREFLPLEVYTRAEALADTGLDY 731
Qy 676 ETMFREYNYAQRVSGRLKDLDLSTNRNQFVDAFGSLMDDLGLAGVQGVTVNAVSVATLF 735
Db 732 SEIQRNQLHELRFYDI--DRVYKTDGNMAIMRGLANFPQGLGAVQGVTVLGAAGAA 789
Qy 736 SSIIVGTFINFKPFGGMLIIVGLVFLYFLTKTKIVETAPIKMIYPEIDK--LKER 794
Db 790 LSTVSGIASFIANPFGALATGLVLAAGLVAALAYRYSRUSNPNKALYPIITRALKDD 849
Qy 795 ECKSEIAPISEE-----LER---IVLAWHIHQNSHMETKTRDKPKDS--ILTR--AQN 842
Db 850 PGNRR--PGEEREEFDAKLEQAREMIKYSLSVSAVERQEHKAKSKAARLLATRLTQL 907
Qy 843 MLRKS--GYSNLKNAE 857
Db 908 ALRRRAPPEYQQLPMAD 924

```

RESULT 36 RBP2_PLAVB

```

ID RBP2_PLAVB STANDARD; PRT; 1251 AA.
AC Q00799;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE Reticulocyte binding protein 2 (Fragment).
GN RBP2.
OS Plasmodium vivax (strain Belem).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=31273;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92315338; PubMed=1617731;
RA Galinski M.R., Medina C.C., Ingravallo P., Barnwell J.W.;
RT "A reticulocyte-binding protein complex of Plasmodium vivax
merozoites.";
RL Cell 69:1213-1226(1992).
CC -!- FUNCTION: INVOLVED IN RETICULOCYTE ADHESION. SPECIFICALLY BINDS TO
HUMAN RETICULOCYTE CELLS.

```

```

CC -!- SUBCELLULAR LOCATION: Membrane-bound (Probable).
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL; M88098; AAA29744.1; -.
KW Malaria; Receptor; Membrane.
FT NON TER 1
FT NON TER 1251
SQ SEQUENCE 1251 AA; 143741 MW; 54BA51C7404AC572 CRC64;

Query Match 3.2%; Score 144; DB 1; Length 1251;
Best Local Similarity 19.0%; Pred. No. 0.46;
Matches 165; Conservative 138; Mismatches 319; Indels 246; Gaps 42;

Qy 109 SDMVHSGILLIYKONI-IPMPRV-----RKVRKVVT---TSTV--YNGIYSDSITN 155
Db 506 STINIEGALKESGNVEIGFLEKLEIGKVRKLVKVDITKKSINSTVGNFSSLFNFPDLN 565
Qy 156 QHTFYKSIETPMEKMDTIYQCF-----NSLRNTGGNLLTYVDRDDIN 199
Db 566 QYDENKNINDYE-NKMGIEYNEFGSLNKISENLNASENTSDYNSAKTLRLAQKEKVN 624
Qy 200 MTFVLPQVDGTPVVKYGS-----QPELYEPGFWFGSVRRRTTV 240
Db 625 LLNKEEEANKYLRDVKVKSFRFIENFKESLDKINEMIKKEQLTVNEG--HGNVKQLVEN 682
Qy 241 NCELMDMFARNPPDFPVATGDTVMSPPWSEGDDEH--KMEKPFWFSVINNYKVD 299
Db 683 IKELVD-----ENLUSDILKQATGKNEIQKI-----THSTLKNKAKTILGHVDTSAKVVG 733
Qy 300 YQNRG-----TVPLGKTRI-----FLDREBYTLISWEKHLKMSYCPILTMKAFYNGIQT 348
Db 734 IKITPELALTELLGDAKLKTAQELKFESKNNVLETNMKNKTVELDV-----HKNIQD 787
Qy 349 EHSYSYHVFANDITASFTTSKEDMKEF-----NTTY--HCLNE-----EIKABIEKYYA 395
Db 788 AYKVALEILAH--SDEIDTKQKSSKLIEMGNQIYLVKVLINQYKNKISSIKSEEAHSV 845
Qy 396 KVNSTHSGYGLKYF-----KTDGGLYLVMOPLQ-----NRLLDKKNLNKNETYSRRRQ 447
Db 846 KIGNVSKHSLSKITCSDKSYNIIALEKQTELQNLURNSTQETKWTNSDKLEKIKTD 905
Qy 448 AESTTDPMMETNGAGGEYSSENSITVAQVAYDNLRIINNILEDLSKAWCREQHRA 507
Db 906 FESLKNALKTLEGEVNAKASSDN-----HEH-- 932
Qy 508 ALVWNLKINPTSMYINRPVSARKIGDVISVNCIVVDQTSV--SLHKSLRLLSASD 566
Db 933 --VQSKSEPVNP-----ALSE--IEKEETDIDSLTALDELKKG 968
Qy 567 EKC--FSRPPVTFKEMDSTIYKQLGVNNEILLTTTLYLETCOENTYEVYFOAKTDMYIKN 625
Db 969 RTCEVSR---YKLKIDT-----VTKEISDDTELINTIEKNVAYL-----AYIKKN. 1011
Qy 626 Y-----BHLKTVPLS-----SITLDTFIALNFTLLENVDFKVIETY 662
Db 1012 YEDTVQDVLTINEHFNFKQVSNHEPTNFDKSNKSSEELTKAVTDSKTIISKLGKVIIEVN 1071
Qy 663 TRDEKRL--SNVFDIETMFREYNYAQRVSGLRK-----DLDDLTNRNQFVDA-- 709
Db 1072 ENTEMNTIESSAKEIEAL---YNELKNKTSLSNEIYQTSNEVKLQEMKSNADKVIDVSKI 1128
Qy 710 FGSLMDLGLAGVQGVTVNAVSGVATLFSISYVTG--FINFKNPPFGGMLIIVGLVFAIY 767
Db 1129 FNTVLD-----TQKSNIVTNOHNSNNVKNVKKLQKLELIDSSFTLE 1171
Qy 768 FLTCKTKIYETAPIKMIYPEIDKLEREGKSEIAPISEEELERIVLAWHIHQNSHMETK 827

```

```

Db 1172 SIKKNEIY--SHIKTNGIEQL-QQTNKSEHNVAKHK-EKIV---HLINRVESLKG 1224
QY 828 TRKDPKDSILTRAQMLRKSGYGNLKN 855
Db 1225 VKNHDDDOYMKKLNASLN---DNIKN 1248

RESULT 37
KIPI_YEAST
ID_KIPI_YEAST STANDARD; PRT; 1111 AA.
AC P28742;
DT 01-DEC-1992 (Rel. 24, Created)
DT 01-DEC-1992 (Rel. 24, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE Kinesin-like protein KIPI.
GN KIPI OR CIN9 OR YBL063W OR YBL0504 OR YBL0521.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=S288C;
RX MEDLINE=92317166; PubMed=1618910;
RA Roof D.M., Meluh P.B., Rose M.D.;
RT "Kinesin-related proteins required for assembly of the mitotic
spindle.";
RL J. Cell Biol. 118:95-108(1992).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=S288C;
RX MEDLINE=94205266; PubMed=8154187;
RA Scherens B., el Bakoury M., Vierendeels F., Dubois E., Messenguy F.;
RT "Sequencing and functional analysis of a 32,560 bp segment on the
left arm of yeast chromosome II. Identification of 26 open reading
frames, including the KIPI and SEC17 genes.";
RL Yeast 9:1355-1371(1993).
RN [3]
RP CHARACTERIZATION.
RC STRAIN=S288C;
RX MEDLINE=92354062; PubMed=1643659;
RA Saunders W.S., Hoyt M.A.;
RT "Kinesin-related proteins required for structural integrity of the
mitotic spindle.";
RL Cell 70:451-458(1992).
CC -!- FUNCTION: REQUIRED FOR ASSEMBLY OF THE MITOTIC SPINDLE. INTERACT
WITH SPINDLE MICROTUBULES TO PRODUCE AN OUTWARDLY DIRECTED
FORCE ACTING UPON THE POLES. FOLLOWING SPINDLE ASSEMBLY, CIN8 AND
KIPI APPARENTLY ACT TO OPPOSE A FORCE THAT DRAWS SEPARATED POLES
BACK TOGETHER. THIS FORCE SEEMS TO BE MEDIATE BY KAR3.
CC -!- SUBUNIT: MIGHT BE DIMERIC.
CC -!- SUBCELLULAR LOCATION: SPINDLE MICROTUBULES THAT LIE BETWEEN THE
POLES.
CC -!- SIMILARITY: BELONGS TO THE KINESIN-LIKE PROTEIN FAMILY. BIMC
SUBFAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC
CC EMBL; Z11962; CAA78019.1; -
CC EMBL; Z23261; CAA80785.1; -
CC EMBL; Z35824; CAA84883.1; -
CC PIR; A42640; A42640.
CC HSP; P17119; 3KAR.
CC SGD; S0000159; KIPI.
CC GO; GO:0005816; C:spindle pole body; IDA.
CC GO; GO:0005200; F:structural constituent of cytoskeleton; IPI.

```

```

DR GO; GO:0007020; P:microtubule nucleation; IPI.
DR GO; GO:000092; P:mitotic anaphase B; IMP.
DR InterPro; IPR001752; kinesin_motor.
DR Pfam; PF00225; kinesin; 1.
DR PRINTS; PRO0380; KINESINHEAVY.
DR SMART; SM00129; KISC; 1.
DR PROSITE; PS00411; KINESIN MOTOR DOMAIN1; 1.
DR PROSITE; PS0067; KINESIN MOTOR DOMAIN2; 1.
KW Motor protein; Cell division; Microtubules; ATP-binding; Coiled coil;
KW Mitosis; Cell cycle.
FT DOMAIN 50 423 KINESIN-MOTOR (BY SIMILARITY).
FT DOMAIN 424 510 COILED COIL (POTENTIAL).
FT DOMAIN 648 670 COILED COIL (POTENTIAL).
FT DOMAIN 710 780 COILED COIL (POTENTIAL).
FT DOMAIN 808 828 COILED COIL (POTENTIAL).
FT NP_BIND 141 148 ATP (BY SIMILARITY).
SQ SEQUENCE 1111 AA; 125794 MW; 212F8279766137FC CRC64;

Query Match 3.1%; Score 140; DB 1; Length 1111;
Best Local Similarity 21.1%; Pred. No. 0.69;
Matches 121; Conservative 98; Mismatches 199; Indels 156; Gaps 29;

QY 232 GSYRRRTTNCCELMDFARNPDPFFVTATGDTVMSPPWSGEDDHENKMKHKMFVSV 291
Db 264 GSLKRKVAAT-KCNDLSSRSHTVF-----TITTNIVEQ-----DSKDHGQNKNFVK----- 308
QY 292 INNKYKVDYQNRGTVPGLKTRIFLDREEVTLSEKHLKMSYCPILTLWKAFYNGIQTEHS 351
Db 309 IGKLNVLVDL-----AGSENINRSGAENKRAQEAGLINKSL--LTGLRVI--NAL--VDHS 357
QY 352 GSYHFVANDIT-----ASFTTSKEDMKFNTTYHCLNEIEKAEIKYA 395
Db 358 NHIPRESKLTLLQDSLGMYKTCIIATISPAKISMEETAST---LEVATRAKSIKNTP 414
QY 396 KVNSTHSKYGDLYKFYKTDGGLVLMQPLIQNRLLDAKNLNNTYERRRRRQAEESTDPM 455
Db 415 QVNQSLSKDTCCLKDY-----IQBIEKLRNDLKNSRNKQGFITODQL 456
QY 456 MMTGNGAGGEYSSSENSITVAQVOYAYDNLR-----IRNNILEDLSKAWCR 502
Db 457 -----DLYESNILDEQNKLHNLREQIKFKENYLNQLDINNLLQS-----E 500
QY 503 EQHRAALVWN-----ELSKINPTSVMSG---IYNRPVSAKRIGDVISVSNICIVVD 549
Db 501 KEKLIAIQNFVDFSNFYSEIQKIHTNLNLELMNEVIQORDFSLENSQKYNTNQNMQL- 559
QY 550 QTSVSLHLSRLLSAS-----DEKC-----FSRPPVTFKFMNDSTIYKQQLGVNNE 595
Db 560 KISQQVLQTLNTLQGLSNLYNSKCSSEVIKGVTEELTRNVNTHKAKHDSL-KSLNLTNN 618
QY 596 ILIT-----TTYLETQ-ENTYFYQAKTDMY-----IYKNYEH-----LKTVPIS 635
Db 619 LLMNMNELVRISLTSLEIFQSDSTSHYRKDLNEIYQSHQQFLKNTQNDIKSKCLDSIGSS 678
QY 636 STITLDTFTIALNFT-----LLENV---DFKVIETYRDEKRLSNVFDIETMFREYN 684
Db 679 ILTSINE-ISOCTTNLSNMVLIENQSQSGSKLKEQDLKIKLKN--DLINERRISNQ 735
QY 685 YAQRVSGLRKDLDD-LSTNRNQFVDFAGSLMDDL 717
Db 736 FNQQLAEMKRYFQDHVSRTRSEFHDNLKNCIDNL 769

RESULT 38
CUT7_SCHPO
ID CUT7_SCHPO STANDARD; PRT; 1085 AA.
AC P24339;
DT 01-MAR-1992 (Rel. 21, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Kinesin-like protein cut7.
GN CUT7 OR SPAC25G10.07C.
OS Schizosaccharomyces pombe (Fission yeast).

```

OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;
 OC Schizosaccharomycetes.
 OX NCBI_TaxID=4896;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=91015362; PubMed=2145514;
 RA Hagan I., Yanagida M.;
 RT "Novel potential mitotic motor protein encoded by the fission yeast
 cut+ gene."
 RL Nature 347:563-566(1990).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=972;
 RA Wood V., Williams R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
 RA Sgouras J., Peat N., Hayes J., Baker S., Basmah D., Bowman S.,
 RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
 RA Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,
 RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
 RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
 RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
 RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,
 RA Oliver K., O'Neil S., Pearson D., Quail M.A., Rabinowitsch E.,
 RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
 RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
 RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
 RA Woodward J., Volckaert G., Aert R., Robben J., Grynoprez B.,
 RA Weljens J., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,
 RA Gabel C., Fuchs M., Fritz C., Holzer E., Moestl D., Hilbert H.,
 RA Borzym K., Langer I., Beck A., Lehnach H., Reinhardt R., Pohl T.M.,
 RA Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,
 RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,
 RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
 RA Lucas M., Rochet M., Gallard C., Tallada V.A., Carlson A., Thode G.,
 RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
 RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
 RA Cerutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
 RA Shpakovski G.V., Ussery D., Barrell B.G., Nurse P.;
 RT "The genome sequence of Schizosaccharomycetes pombe";
 RL Nature 415:871-880(2002).
 CC -!- FUNCTION: COULD BE A SPINDLE POLE BODY MOTOR. ON TRANSITION FROM
 CC G2 TO M PHASE OF THE CELL CYCLE, THE SPINDLE POLE BODY DUPLICATES;
 CC THE DAUGHTER POLE BODIES SEED MICROTUBULES WHICH INTERDIGITATE TO
 CC FORM A SHORT SPINDLE THAT ELONGATES TO SPAN THE NUCLEUS AT
 CC METAPHASE. MUTATIONS AT CUT7 BLOCK SPINDLE FORMATION.
 CC -!- SIMILARITY: BELONGS TO THE KINESIN-LIKE PROTEIN FAMILY. BIMC
 CC SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 CC EMBL; X57513; CAA04738.1; --
 CC EMBL; Z70691; CAA94636.1; --
 CC FIR; T38378; T38378.
 CC HSP; P17119; 3KAR.
 CC GeneDB Spombe; SPAC25G10.07c; --
 CC InterPro; IP0001752; kinesin_motor.
 CC Pfam; PF00225; kinesin; 1.
 CC PRINTS; PR00380; KINESINHEAVY.
 CC SMART; SM00129; KISC; 1.
 CC PROSITE; PS00411; KINESIN MOTOR DOMAIN1; 1.
 CC PROSITE; PS00417; KINESIN MOTOR DOMAIN2; 1.
 CC Motor protein; Cell division; Microtubules; ATP-binding; Coiled coil;
 KW Mitosis; Cell cycle; Phosphorylation; Repeat.
 KW DOMAIN 70 435 KINESIN-MOTOR (BY SIMILARITY).
 FT DOMAIN 436 604 COILED COIL (POTENTIAL).
 FT DOMAIN 715 740 COILED COIL (POTENTIAL).
 FT

FT DOMAIN 897 955 COILED COIL (POTENTIAL).
 FT NP_BIND 159 166 ATP (BY SIMILARITY).
 FT REPEAT 987 998
 FT REPEAT 999 1010
 FT MOD_RES 1011 1011 PHOSPHORYLATION (BY CDC2) (BY
 FT CONFLICT 34 61 SIMILARITY).
 FT SASNPKRREPPPTDTGYDPRSDTNSPT ->
 FT LRAILGNDSLLLTLL (IN REF. 1).
 SQ SEQUENCE 1085 AA; 122133 MW; 5669277875559D58 CRC64;
 Query Match 3.1%; Score 139; DB 1; Length 1085;
 Best Local Similarity 19.9%; Pred.No. 0.77;
 Matches 164; Conservative 106; Mismatches 311; Indels 244; Gaps 41;
 QY 152 SITNOHFYKIGIEPWETEKMDTIYQCFNSLRNLNTGNGLLTYVDRDDINMTVFLQPVDCVT 211
 DB 26 STPNSH-FRSASNPKRREPPPTI-----DTG-----YDPSDVTN-----SPTDHAL 65
 QY 212 PDVKRYGSGPELYLEPGFMGYSRRRTTVNCEL-----MDMPARSN----- 252
 DB 66 HD-----ENETINVVVRGRTDQEVDRDNSSSLAVSTSGAMGAE 104
 QY 253 -----PPDPFVTATGDTVEMSPFWSGDDH-----ENKMHKPKWFSVNNYK--VVDYQ 301
 DB 105 LAIOSDFSSMLVT---KTYAFDKVGPEDADQLMFLFNSV--APMLEQVLNGYNCTIFAYG 159
 QY 302 NRGTVPLGKTRIFLDREBYTILSWEKHLKNMSYCPILTWKAFYNGIQTSHSGSYHFVANDI 361
 DB 160 QTGT---GKT-----YTMGSD-----LSDSDGILSEGAGLIPRALYQL 194
 QY 362 TASPTTKEDMKFNTYTHCLNEEIK---AEIEKKYAKVNSTHSHKYGD-----KYF 410
 DB 195 FSSLDNSNQEVAVKSYVELNNEIRDLLYSEELKPARVFEDTSRRGNVITGIEEYI 254
 QY 411 KTDG-GLYLVWQPLIQNRLLDAK-NKLNNEYTS-----RRSRQAESTTDPMMWMTGNAG 464
 DB 255 KNAGDGLRLREGSHRRQVAATKCNLDLSSRSHTFTTLHRKVSSTGMDTNSLTNNNS 314
 QY 465 GEYSSENSITVAQV-----QYAYDNLRT-----INNILEDLSK---AWCREOHRALV 510
 DB 315 DDLRLASKLHVDLAGSENIGRGAENKRARETGMNQSLTLGRVINALVEKAHPIYR 374
 QY 511 WNELSKINPTSVMSIMYNPVPSAKRIGDIVSVSNCI VVDQTSVSLHLSRLLS-ASDEKC 569
 DB 375 ESKLTRLQDS-----LGKTKTSMIVTVSTNTNLEETISTILEVAARAKS 420
 QY 570 FSRPPTVTFKFMNDSTIY-----KQQLGV-----NNEILLTTTYLE----- 604
 DB 421 IRNKP-----QNNQLVFRKVLIKDLVLDIERLKNLKNATKKNGVYLAESTYKELMDRVQ 475
 QY 605 -----TCOENTEY-----FOAKTD-----MYIYK-NYEHKTVPLSSITTLDTFIALNFTLL 651
 DB 476 NKDLLCOFQARKLEVLNKNVKSREQLQYVSKSQEHKKEVEALQQLVNSSTESLVKS 535
 QY 652 ENVDFFK---VIEL-----YTRDEKRLSNV-FDIETMFE-YNYVAQVSGRLKDLDLST 701
 DB 536 ENEKKNELVLEIEKRKKYETNEAKITVTATLSQYRESKEYIASLYEKLDRTERNNKE 595
 QY 702 NRNOFVDAFGSLMDDDLGAGQTVVNAVSGVATL---FSSIVTGFINFIPKPFQGMGLMIIV 758
 DB 596 NENNFNLKFNLLTMLRSFHGSDFTDNGYFTLLDNFNASMEELLNTHSN----- 645
 QY 759 VIGVLPALYFLTKTKIYETAPIKMIYPEIDKLKEREGKSEIAPISSEELERIVLANHIH 818
 DB 646 -----QLLISMKTITEH-----FQSLDEALQSARSSCAVPNS--SLDLIVSEL--- 686
 QY 819 QONSHMETKTRKDPKDSILTRAQNMRLKRSGYS-NLKNAESVEML 862
 DB 687 -----KOSKNSLLDALEHSLQDISSWSOKJGNGISSLEI 720
 RESULT 39
 SYFB_MYCGE

Proc. Natl. Acad. Sci. U.S.A. 90:11172-11176(1993).

[2] SEQUENCE FROM N.A.

RA Viessers S., Urtestarazu L.A., Jauniaux J.-C.;

RL Submitted (MAR-1994) to the EMBL/GenBank/DBJ databases.

[3] SEQUENCE OF 1-3457 FROM N.A.

RP MEDLINE=94052110; PubMed=8234262;

RA Li Y.-Y., Yeh E.-Y., Hays T., Bloom K.S.;

RT "Disruption of mitotic spindle orientation in a yeast dynein mutant.";

RL Proc. Natl. Acad. Sci. U.S.A. 90:10096-10100(1993).

CC -!- FUNCTION: Cytoplasmic dynein acts as a motor for the intracellular

CC retrograde motility of vesicles and organelles along microtubules.

CC Dynein has ATPase activity; the force-producing power stroke is

CC thought to occur on release of ADP. Required to maintain uniform

CC nuclear distribution in hyphae. May play an important role in the

CC proper orientation of the mitotic spindle into the budding

CC daughter cell yeast. Probably required for normal progression of

CC the cell cycle.

CC -!- SUBUNIT: Consists of at least two heavy chains and a number of

CC intermediate and light chains.

CC -!- SUBCELLULAR LOCATION: Cytoplasmic. Probably binds indirectly to

CC the inner plasma membrane.

CC -!- DOMAIN: Dynein heavy chains probably consist of an N-terminal stem

CC (which binds cargo and interacts with other dynein components),

CC and the head or motor domain. The motor contains six tandemly-

CC linked AAA domains in the head, which form a ring. A stalk-like

CC structure (formed by two of the coiled coil domains) protrudes

CC between AAA 4 and AAA 5 and terminates in a microtubule-binding

CC site. A seventh domain may also contribute to this ring; it is not

CC clear whether the N-terminus or the C-terminus forms this extra

CC domain. There are four well-conserved and two non-conserved ATPase

CC sites, one per AAA domain. Probably only one of these (within AAA

CC 1) actually hydrolyzes ATP, the others may serve a regulatory

CC function.

CC -!- SIMILARITY: Belongs to the dynein heavy chain family.

CC

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch).

CC

DR EMBL; Z21977; CAAT9923.1; -

DR EMBL; L15626; AAAL6055.1; -

DR EMBL; Z28279; CAAB2132.1; -

DR PIR; S38128; S38128.

DR SGD; S0001762; DYN1.

DR GO; GO:0005881; C:cytoplasmic microtubule; IDA.

DR GO; GO:000132; P:mitotic spindle orientation; IGI.

DR InterPro; IPR003593; AAA ATPase.

DR InterPro; IPR004273; Dynein heavy.

DR Pfam; PF03028; Dynein heavy; 1.

DR SMART; SM00382; AAA; 3.

DR Karyogamy; Motor protein; Dynein; Microtubules; ATP-binding; Repeat;

KW Coiled coil.

FT DOMAIN 1 1757 STEM (BY SIMILARITY).

FT DOMAIN 1758 1979 AAA 1 (BY SIMILARITY).

FT DOMAIN 2036 2273 AAA 2 (BY SIMILARITY).

FT DOMAIN 2379 2628 AAA 3 (BY SIMILARITY).

FT DOMAIN 2722 2984 AAA 4 (BY SIMILARITY).

FT DOMAIN 2993 3300 STALK (BY SIMILARITY).

FT DOMAIN 3370 3599 AAA 5 (BY SIMILARITY).

FT DOMAIN 3760 3970 AAA 6 (BY SIMILARITY).

FT DOMAIN 154 175 COILED COIL (POTENTIAL).

FT DOMAIN 486 508 COILED COIL (POTENTIAL).

FT DOMAIN 542 566 COILED COIL (POTENTIAL).

FT DOMAIN 932 959 COILED COIL (POTENTIAL).

FT DOMAIN 1042 1063 COILED COIL (POTENTIAL).

FT DOMAIN 1681 1705 COILED COIL (POTENTIAL).

FT DOMAIN 2993 3092 COILED COIL (POTENTIAL).

FT	DOMAIN	3242	3300	COILED COIL (POTENTIAL).
FT	DOMAIN	3532	3608	COILED COIL (POTENTIAL).
FT	NP_BIND	1796	1803	ATP (POTENTIAL).
FT	NP_BIND	2074	2081	ATP (POTENTIAL).
FT	NP_BIND	2418	2425	ATP (POTENTIAL).
FT	NP_BIND	2760	2767	ATP (POTENTIAL).
FT	CONFLICT	589	589	Y -> C (IN REF. 3).
FT	CONFLICT	601	601	V -> A (IN REF. 3).
FT	CONFLICT	1364	1364	E -> A (IN REF. 3).
FT	CONFLICT	2118	2119	ML -> IV (IN REF. 1).
SQ	SEQUENCE	4092	AA; 471337	MM; 3D9DF447E8E2D6BB CRC64;

Query Match 3.0%; Score 134.5; DB 1; Length 4092;

Best Local Similarity 19.9%; Pred. No. 9.1;

Matches 119; Conservative 103; Mismatches 234; Indels 141; Gaps 28;

QY 336 LTLKAFKAFNGIQTEHS---GSYHFVANDITASFTTSKEDMKEFNTTYHCLNEEIKAEIEK 392

DB 441 LSTRKRFFSLAETIKSISPTSYHEEIQRLYHPF-----EQIHDIISVNFRLKLEQAESEFSK 496

QY 393 KY-----AKVNSTHSGYGLKFKTDGGLV-LV-WOPL-----IQNRLDAAKNNLNETYS 441

DB 497 NMLDLERKLQNTLASFMDSDHCPTEKUSLVKFKPLMELCPRIKVKVLENQOILLEL-- 554

QY 442 RRSRRQAEISTDPMMEMTGAGGEYSSENSITVAQVQAYD-NLRIRINNILEDLKAW 500

DB 555 KKDRIQLTGLLELPKIL-----HVEALNIPISARISYFLNVQSRIDNIV----- 601

QY 501 CREQRAALV---WNELSKINPTSVMSIYNRPVSARKRIGDIVSVNSCIVVDQTSVSLHK 557

DB 602 ---QVLEALFGSNWD-----TLEGR-----SISTSIQVLRKETNPHD 636

QY 558 -SLRLLSASDKCFSR---PPVTFKFMNDSTIYKGLGVNNEILLTITTYETCQENTEY 613

DB 637 VFLWLGNFPEKATANLLTTPILKLRNEDDY--ELKVNDFDLAAAYSE-LRSLTYMA 693

QY 614 FQ-----AKTDWYIKVNVHLKTVPLSSITLDTFI---ALNFTLENVDFK---- 657

DB 694 FQVPSHVIRIARTWMLYPR-----AINVELIQTFSSLSKLSYTFYNIPLKRNQV 746

QY 658 ---VIE--LYTRDEKLSNVFDIETMREYNYAQRVSGLRKOLLDTSTNRNQVDAFG 711

DB 747 TWMLLQILITPWSLEQESSEMSCSVHSLARLEKSIDGILSDYQILKNSEPOPAKEFS 806

QY 712 SI-----MDDLGAQGVTVNAVSGVATLFSIVTGFNFINKPFGGMLMIIIVIGVLP 765

DB 807 GLKSGFDGTADDLHEVEEIIISNIAIFENLFTKGLTNVSDHI-STFNLLIISILEKVLN 865

QY 766 IYFLTTKTKIYETAPIKMIYPE-IDKLKEREGKSEIAP-----SEELERIVL 813

DB 866 LK-----KMHFPKGVKLKLSFNEGRITSSPSLAAMKRSLLKDLKLEALLNKVVL 911

QY 814 AMHIHQNSHMETKTR-----KDPKDSILTBAQNMLRKRSYSLNKKNAESVEML 862

DB 912 INFLHDPDPLSTLTITFNSLVILKDKDQNCIEQVQNLHCKINSY--VKEQKMEFL 966

RESULT 41

MLP1 YEAST

ID MLP1 YEAST STANDARD; PRT; 1875 AA.

AC Q02455;

DT 01-OCT-1993 (Rel. 27, Created)

DT 01-JUN-1994 (Rel. 29, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Myosin-like protein MLP1.

GN MLP1 OR YKR095W OR YKR415.

OS Saccharomyces cerevisiae (Baker's yeast).

OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;

OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.

OX NCBI_TaxID=4932;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=S288c;


```

RESULT 43
BXD_CLOBO
ID BXD_CLOBO STANDARD; PRT; 1276 AA.
AC P19321;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Botulinum neurotoxin type D precursor (EC 3.4.24.69) (BONT/D)
DE (Bontoxilysin D).
GN BOTD.
OS Clostridium botulinum.
OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC Clostridium.
OX NCBI_TaxID=1491;
RN [1]_
RP SEQUENCE FROM N.A.
RC STRAIN=BVD/-3;
RC MEDLINE=91016853; PubMed=2216736;
RX Binz T., Kurazono H., Popoff M.R., Eklund M.W., Sakaguchi G.,
RA Kozaki S., Krieglstein K., Henschen A., Gill D.M., Niemann H.;
RA "Nucleotide sequence of the gene encoding Clostridium botulinum
RT neurotoxin type D.";
RL Nucleic Acids Res. 18:5556-5556(1990).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CB16;
RC MEDLINE=93042276; PubMed=1420572;
RX Sunagawa H., Ohyama T., Watanabe T., Inoue K.;
RA "The complete amino acid sequence of the Clostridium botulinum type D
RT neurotoxin, deduced by nucleotide sequence analysis of the encoding
RT phase d-16 phi genome.";
RL J. Vet. Med. Sci. 54:905-913(1992).
RN [3]
RP PARTIAL SEQUENCE.
RC STRAIN=D-SA, and D-1873;
RC MEDLINE=89139741; PubMed=2668193;
RX Moriishi K., Syuto B., Kubo S., Oguma K.;
RA "Molecular diversity of neurotoxins from Clostridium botulinum type D
RT strains.";
RL Infect. Immun. 57:2886-2891(1989).
RN [4]
RP IDENTIFICATION OF SUBSTRATE.
RX MEDLINE=94230352; PubMed=8175689;
RX Yamaeaki S., Baumeister A., Binz T., Blasi J., Link E., Cornille F.,
RA Roques B., Fyксе E.M., Suedhof T.C., Jahn R., Niemann H.;
RA "Cleavage of members of the synaptobrevin/VAMP family by types D and
RT F botulin neurotoxins and tetanus toxin.";
RL J. Biol. Chem. 269:12764-12772(1994).
CC -!- FUNCTION: BOTULINUS TOXIN ACTS BY INHIBITING NEUROTRANSMITTER
CC RELEASE. IT BINDS TO PERIPHERAL NEURAL SYNAPSES. IS INTERNALIZED

```


Qy 504 OHRAALVWNELSKINPTSVMSMIYNRPVSAKRIGDVISVNCIWDOTSVSLHKSLELLS 563
Db 1442 -----YIYEKKYLOHCLLKHNDYKVELPRKD-INYSQMESIKTRNFH-----S 1489
Qy 564 ASDEKCFSPRPVTFKFW-----NDSTIYKGOLGVNNEILLTTTYLETQCENTEY--- 612
Db 1490 LSEOFASFNLILSFYIIKNDNDNVYKNQYIYNKYIYNKNSICNKNVI--CNKNYIYNKN 1547
Qy 613 YFOAKTDMYIYKN-YEHLKTVPLS 635
Db 1548 NIYNKNYIYNKNILTHAKSVLLS 1571

Search completed: October 28, 2003, 15:31:34
Job time : 38 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 28, 2003, 15:27:11 ; Search time 118 Seconds
(without alignments)
1891.657 Million cell updates/sec

Title: US-10-055-364-24
Perfect score: 4547
Sequence: 1 MAGSLKRGSLALWLYQV.....KRGYSNLKNAESVEMLNTL 865

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Maximum Match 100%
Listing first 45 summaries

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_23.*
1: sp_archaea.*
2: sp_bacteria.*
3: sp_fungi.*
4: sp_human.*
5: sp_invertebrate.*
6: sp_mammal.*
7: sp_mhc.*
8: sp_organelle.*
9: sp_phage.*
10: sp_plant.*
11: sp_rodent.*
12: sp_virus.*
13: sp_vertebrate.*
14: sp_unclassified.*
15: sp_rvirus.*
16: sp_bacteriap.*
17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	ID	Description
1	4547	100.0	Q913V4	Q913V4 porcine gam
2	4547	100.0	Q8JYD3	Q8JYD3 porcine lym
3	4271	93.9	Q8B3X4	Q8B3X4 porcine lym
4	4265	91.8	Q8B423	Q8B423 porcine lym
5	3529	77.6	Q8AYX3	Q8AYX3 porcine lym
6	2513	55.3	Q91A92	Q91A92 porcine lym
7	2504	55.1	Q91A90	Q91A90 porcine herpe
8	2493.5	54.8	Q91A89	Q91A89 porcine herpe
9	2491	54.8	Q91A91	Q91A91 porcine herpe
10	2489.5	54.8	Q91A91	Q91A91 porcine herpe
11	1962.5	43.2	Q8JUX2	Q8JUX2 badger herp
12	1953.5	43.0	Q86613	Q86613 equine herp
13	1937	42.6	Q88906	Q88906 caposi's sa
14	1927.5	42.4	Q9WR79	Q9WR79 macaca mula
15	1927.5	42.4	Q9J2M4	Q9J2M4 thesus monk
16	1925.5	42.3	Q9J2M4	Q9J2M4 macaca mula

RESULT 1

Q913V4	PRELIMINARY;	PRT;	865 AA.
AC	Q913V4:		
DT	01-DEC-2001 (Tremblrel. 19, Created)		
DT	01-DEC-2001 (Tremblrel. 19, Last sequence update)		
DT	01-MAR-2002 (Tremblrel. 20, Last annotation update)		
DE	Envelope glycoprotein B.		
GN	GPB.		
OS	Porcine gamma-herpesvirus.		
OC	Viruses; dsDNA viruses, no RNA stage; Herpesviridae;		
OC	Gammaherpesvirinae.		
OX	NCBI TaxID=169968;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RX	MEDLINE=2112435;		
RA	Huang C., Fuchimoto Y., Gleit Z., Ericsson T., Greissner A.,		
RA	Schier-Dolberg R., Melendy E., Kitamura H., Fishman J., Ferry J.,		
RA	Harris N., Patience C., Sachs D.;		
RT	"Posttransplantation lymphoproliferative disease in miniature swine		
RT	after allogeneic hematopoietic cell transplantation: similarity to		
RT	human PTLD and association with a porcine gammaherpesvirus.";		
RL	Blood 97:1467-1473(2001).		
DR	EMBL; AY044427; AAK95853.1; -		
DR	InterPro; IPR000234; Glycoprot_B.		
DR	Pfam; PF00606; Glycoprotein_B; 1.		
DR	ProDom; PD000693; Glycoprot_B; 1.		
SQ	SEQUENCE 865 AA; 99291 MW; E3F6CEFI1B3EB326E CRC64;		
Query Match 100.0%; Score 4547; DB 12; Length 865;			
Best Local Similarity 100.0%; Pred. NO. 1.1e-277;			
Matches 865; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
Qy	1	MAGSLKRGSLALWLYQVSLTAETGVTSPPNTATWSTESPLTGHYTHDSHGE	60
Db	1	MAGSLKRGSLALWLYQVSLTAETGVTSPPNTATWSTESPLTGHYTHDSHGE	60
Qy	61	RGNNENRDSSEQNKNVGSPTFPYRVCASGVGDVFRFOTDHYCPDASDMVHSEGI	120

```

Db      61  RGNENRDESEONKNIYGSPTFPYRVCSASGVGVDFRQTDHVCPCDASDMVHSEGIILLI 120
Qy      121  YKQNIIPFMRVRKRVKVVTTSTVYNGIYSDSITNQHTFYKSTBPWETEKMDTIYQCFNS 180
Db      121  YKQNIIPFMRVRKRVKVVTTSTVYNGIYSDSITNQHTFYKSTBPWETEKMDTIYQCFNS 180
Qy      181  LRLTNGNLLTYVDRDDINMTVFLQPDVGVTPDKRYGSOPELYLEPGFWGVSRRRTTV 240
Db      181  LRLTNGNLLTYVDRDDINMTVFLQPDVGVTPDKRYGSOPELYLEPGFWGVSRRRTTV 240
Qy      241  NCELMDMFARSNPPDFVTATGDTVEMSPFWSGEDDHNKMKHEKMPFVSVINNYKVVDY 300
Db      241  NCELMDMFARSNPPDFVTATGDTVEMSPFWSGEDDHNKMKHEKMPFVSVINNYKVVDY 300
Qy      301  QNRGTVPGLKTRIFLDREBYTLSWEKHLKNMSYCPPLTLWKAIFYNGIQTSHSGSYHFVAND 360
Db      301  QNRGTVPGLKTRIFLDREBYTLSWEKHLKNMSYCPPLTLWKAIFYNGIQTSHSGSYHFVAND 360
Qy      361  ITASFTTSKEDMKFNTTYHCLNEEIKAEIEKKYAKVNSTHSGYDGLKYFKTDGGLYLW 420
Db      361  ITASFTTSKEDMKFNTTYHCLNEEIKAEIEKKYAKVNSTHSGYDGLKYFKTDGGLYLW 420
Qy      421  QPLIQNRLLDLDAKNNLNNETYSRRSRQAESTTDPMMEMTNGAGGEYSSENSITVAQVOY 480
Db      421  QPLIQNRLLDLDAKNNLNNETYSRRSRQAESTTDPMMEMTNGAGGEYSSENSITVAQVOY 480
Qy      481  AYDNLRIIRINNILEDLSKAWCREQRAALVMNLSKINPTSVMSMIYNRPVSAKRIGDVI 540
Db      481  AYDNLRIIRINNILEDLSKAWCREQRAALVMNLSKINPTSVMSMIYNRPVSAKRIGDVI 540
Qy      541  SVSNCIYVDQTSVSLHSLRLLSASDEKCFSPRPVTFKFMNDSTIYKQGLGVNNEILLTT 600
Db      541  SVSNCIYVDQTSVSLHSLRLLSASDEKCFSPRPVTFKFMNDSTIYKQGLGVNNEILLTT 600
Qy      601  TVLETQCENTEYFYQAQTDWYIYKNYEHLKTVPLSSITTLDTFIALNFTLLENVDFKVE 660
Db      601  TVLETQCENTEYFYQAQTDWYIYKNYEHLKTVPLSSITTLDTFIALNFTLLENVDFKVE 660
Qy      661  LYTRDEKRLSNVDFIETMFREYNYAQRVSGRLKDLDDLSTNRNQFVDAFGSLMDDLGA 720
Db      661  LYTRDEKRLSNVDFIETMFREYNYAQRVSGRLKDLDDLSTNRNQFVDAFGSLMDDLGA 720
Qy      721  GQTVNAVSGVATLFSIVTGFINFIKNPFGGMLIIVLGVLFAYIYFLTKTKIYETAP 780
Db      721  GQTVNAVSGVATLFSIVTGFINFIKNPFGGMLIIVLGVLFAYIYFLTKTKIYETAP 780
Qy      781  IKMIYPIDKLKREGKSEIAPISEEELERIVLAMHIHQONSHMETKTRDKPKDSILTRA 840
Db      781  IKMIYPIDKLKREGKSEIAPISEEELERIVLAMHIHQONSHMETKTRDKPKDSILTRA 840
Qy      841  QNMLRKRSYGNLKNABSEVEMNLTL 865
Db      841  QNMLRKRSYGNLKNABSEVEMNLTL 865

```

RESULT 2

```

Q8UYD3 ID Q8UYD3 PRELIMINARY; PRT; 876 AA.
AC Q8UYD3
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Glycoprotein B.
OS Porcine lymphotropic herpesvirus 1.
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Gammaherpesvirinae.
OX NCBI_TaxID=91740;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=sample #56;
RX MEDLINE=99226949; PubMed=10211967;
RA Ehlers B., Ulrich S., Goltz M.;
RT "Detection of two novel porcine herpesviruses with high similarity to

```

```

RT gammaherpesviruses.";
RL J. Gen. Virol. 80:971-978(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=sample #56;
RX MEDLINE=20036635; PubMed=10567652;
RA Ulrich S., Goltz M., Ehlers B.;
RT "Characterization of the DNA polymerase loci of the novel porcine
RT lymphotropic herpesviruses 1 and 2 in domestic and feral pigs.";
RL J. Gen. Virol. 80:3199-3205(1999).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=sample #56;
RX MEDLINE=22008485; PubMed=12009880;
RA Goltz M., Ericsson T., Patience C., Huang C.A., Noack S., Sachs D.H.,
RA Ehlers B.;
RT "Sequence analysis of the genome of porcine lymphotropic herpesvirus 1
RT and gene expression during post-transplant lymphoproliferative disease
RT of pigs.";
RL Virology 294:393-393(2002).
DR EMBL; AF478169; AAW22114.1; -.
DR InterPro; IPR000234; Glycoprot_B.
DR Pfam; PF00606; Glycoprotein B; 1.
DR ProDom; PD000693; Glycoprot_B; 1.
SQ SEQUENCE 876 AA; 100705 MW; 2D1D4DB61F5C0D24 CRC64;

```

```

Query Match 100.0%; Score 4547; DB 12; Length 876;
Best Local Similarity 100.0%; Pred. No. 1.le-277;
Matches 865; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy      1  MAGSLKRGSLVALWLYQVALYSIAETVTSPPNTATWSTESPLTGHYTHDSHGE 60
Db      12  MAGSLKRGSLVALWLYQVALYSIAETVTSPPNTATWSTESPLTGHYTHDSHGE 71
Qy      61  RGNENRDESEONKNIYGSPTFPYRVCSASGVGVDFRQTDHVCPCDASDMVHSEGIILLI 120
Db      72  RGNENRDESEONKNIYGSPTFPYRVCSASGVGVDFRQTDHVCPCDASDMVHSEGIILLI 131
Qy      121  YKQNIIPFMRVRKRVKVVTTSTVYNGIYSDSITNQHTFYKSTBPWETEKMDTIYQCFNS 180
Db      132  YKQNIIPFMRVRKRVKVVTTSTVYNGIYSDSITNQHTFYKSTBPWETEKMDTIYQCFNS 191
Qy      181  LRLTNGNLLTYVDRDDINMTVFLQPDVGVTPDKRYGSOPELYLEPGFWGVSRRRTTV 240
Db      192  LRLTNGNLLTYVDRDDINMTVFLQPDVGVTPDKRYGSOPELYLEPGFWGVSRRRTTV 251
Qy      241  NCELMDMFARSNPPDFVTATGDTVEMSPFWSGEDDHNKMKHEKMPFVSVINNYKVVDY 300
Db      252  NCELMDMFARSNPPDFVTATGDTVEMSPFWSGEDDHNKMKHEKMPFVSVINNYKVVDY 311
Qy      301  QNRGTVPGLKTRIFLDREBYTLSWEKHLKNMSYCPPLTLWKAIFYNGIQTSHSGSYHFVAND 360
Db      312  QNRGTVPGLKTRIFLDREBYTLSWEKHLKNMSYCPPLTLWKAIFYNGIQTSHSGSYHFVAND 371
Qy      361  ITASFTTSKEDMKFNTTYHCLNEEIKAEIEKKYAKVNSTHSGYDGLKYFKTDGGLYLW 420
Db      372  ITASFTTSKEDMKFNTTYHCLNEEIKAEIEKKYAKVNSTHSGYDGLKYFKTDGGLYLW 431
Qy      421  QPLIQNRLLDLDAKNNLNNETYSRRSRQAESTTDPMMEMTNGAGGEYSSENSITVAQVOY 480
Db      432  QPLIQNRLLDLDAKNNLNNETYSRRSRQAESTTDPMMEMTNGAGGEYSSENSITVAQVOY 491
Qy      481  AYDNLRIIRINNILEDLSKAWCREQRAALVMNLSKINPTSVMSMIYNRPVSAKRIGDVI 540
Db      492  AYDNLRIIRINNILEDLSKAWCREQRAALVMNLSKINPTSVMSMIYNRPVSAKRIGDVI 551
Qy      541  SVSNCIYVDQTSVSLHSLRLLSASDEKCFSPRPVTFKFMNDSTIYKQGLGVNNEILLTT 600
Db      552  SVSNCIYVDQTSVSLHSLRLLSASDEKCFSPRPVTFKFMNDSTIYKQGLGVNNEILLTT 611
Qy      601  TVLETQCENTEYFYQAQTDWYIYKNYEHLKTVPLSSITTLDTFIALNFTLLENVDFKVE 660
Db      612  TVLETQCENTEYFYQAQTDWYIYKNYEHLKTVPLSSITTLDTFIALNFTLLENVDFKVE 671

```

```
QY 661 LYTRDEKRLSNVFDIETMFREYNYAQRVSGLRKOLLDLSTNRNOFVDAFGLMDDLGA 720
DB 672 LYTRDEKRLSNVFDIETMFREYNYAQRVSGLRKOLLDLSTNRNOFVDAFGLMDDLGA 731
QY 721 GQTVNAVSGVATLSSIVTGFINFKPFGGMLMIIIVIGLVFAIYFLTKTKIYETAP 780
DB 732 GQTVNAVSGVATLSSIVTGFINFKPFGGMLMIIIVIGLVFAIYFLTKTKIYETAP 791
QY 781 IKMIYPEIDKLKEREKSEIPISEELERIVLAMHIHQONSHMETKTRKPKDSILTRA 840
DB 792 IKMIYPEIDKLKEREKSEIPISEELERIVLAMHIHQONSHMETKTRKPKDSILTRA 851
QY 841 QNMLKRKSGYSLKNAESVEMLNTL 865
DB 852 QNMLKRKSGYSLKNAESVEMLNTL 876

RESULT 3
Q8B3X4
ID Q8B3X4 PRELIMINARY; PRT; 876 AA.
AC Q8B3X4;
DT 01-MAR-2003 (TReMBLrel. 23, Created)
DT 01-MAR-2003 (TReMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TReMBLrel. 23, Last annotation update)
DE Glycoprotein B.
OS Porcine lymphotropic herpesvirus 2.
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Gammaherpesvirinae.
OX NCBI_TaxID=91741;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=568;
RA Chmielewicz B., Goltz M., Franz T., Bauer C., Brema S., Ellerbrok H.,
RA Beckmann S., Rziha H.-J., Lahrmann K.-H., Romero C., Ehlers B.;
RT "A novel porcine gammaherpesvirus.";
RL Virology 0:0-0(2003).
DR EMBL; AY170317; AA012353.1; -.
SQ SEQUENCE 876 AA; 100741 MW; 4CB870F431BCB0DF CRC64;

Query Match 93.9%; Score 4271; DB 12; Length 876;
Best Local Similarity 93.4%; Pred. No. 2.5e-260;
Matches 808; Conservative 28; Mismatches 29; Indels 0; Gaps 0;

QY 1 MAGSLKRGSLALWLYQVALYSLSIATGTGVTSPNTATWSTESPLTGHYTHDSSHGE 60
DB 12 MAGSLKRLALALWCQFQVALYSLKAEKATSANTQAWSTEALPIGQYGYDSSHGE 71
QY 61 RGNENRDSSEONKNIYGSPTFPYRVCSAGVGDFRFQTDHVCPCDASDMVHSEGIILLI 120
DB 72 RATSENREDEEHKNIYGSPTFPYRVCSAGVGDFRFQTDHVCPCDASDMVHSEGIILLI 131
QY 121 YKQNIIPFMRVRKRVKVTSTVYNGIYSDSIINQHTFYKSIPEWTEKMDTIYQCFNS 180
DB 132 YKQNIIPFMRVRKRVKVTSTVYNGIYSDSIINQHTFYKSIPEWTEKMDTIYQCFNS 191
QY 181 LRLNTGNNLLTYVDRDDINMTVFLQPDVGTDPVKRYGSOPELYLEPGWFGSYRRRTTV 240
DB 192 LRLNTGNNLLTYVDRDDINMTVFLQPDVGTDPVKRYGSOPELYLEPGWFGSYRRRTTV 251
QY 241 NCELMDFARSNPPDFVTATGTVEMSPWMSGEDDHENKHKPFVSPVNNYKVVDY 300
DB 252 NCELMDFARSNPPDFVTATGTVEMSPWMSGEDDHENKHKPFVSPVNNYKVVDY 311
QY 301 QNRGTVPLGKTRIFLDREYTLWSWKHLKNNMSYCPPLTWKAFYNGIQTEHSGSYHFVAND 360
DB 312 QNRGTVPLGKTRIFLDREYTLWSWKHLKNNMSYCPPLTWKAFYNGIQTEHSGSYHFVAND 371
QY 361 ITASFTTSKEDMKENTYTHCLNEEIKAEIEKKYAKVNSTHYSKYGDLYFKTDGGLYLW 420
DB 372 ITASFTTSKEDMKOPNTYTHCLNEEIKAEIEKKYAKVNSTHYSKYGDLYFKTDGGLFLW 431
QY 421 OPLIQNRLLDKAKNKLNNETYSRRSRQAESTTDDPMWMTGNAGGEYSSENSITVAQVOY 480
```

```
DB 432 QPLIQNRLLDKAKNKLNNETYSRRSRQVDSITGPMMEATGNAGGEYSSENSITVAQVOY 491
QY 481 AYDNLRTIRNNILBEDLSKAWCREQHRALVWNELSKINPTSVMSMIYNRPVSAKRIGDVI 540
DB 492 AYDNLRTIRNNILBEDLSKAWCREQHRALVWNELSKINPTSVMSMIYNRPVSAKRIGDVI 551
QY 541 SVSNICVVDDQTSVSLHKSRLLSASDEKCFRPPVPTFKFMNDSTIYKQGLGVNNEILLTT 600
DB 552 SVSNICVVDDQTSVSLHKSRLLSASDEKCFRPPVPTFKFMNDSTIYKQGLGVNNEILLTT 611
QY 601 TYLETCQENTEYFYQAOKTDMYIYKNEYHLKTVPLSSITTLDTFIALNFTLLENVDFKVE 660
DB 612 TYLETCQENTEYFYQAOKTDMYIYKNEYHLKTVPLSSITTLDTFIALNFTLLENVDFKVE 671
QY 661 LYTRDEKRLSNVFDIETMFREYNYAQRVSGLRKOLLDLSTNRNOFVDAFGLMDDLGA 720
DB 672 LYTRDEKRLSNVFDIETMFREYNYAQRVSGLRKOLLDLSTNRNOFVDAFGLMDDLGA 731
QY 721 GQTVNAVSGVATLSSIVTGFINFKPFGGMLMIIIVIGLVFAIYFLTKTKIYETAP 780
DB 732 GQTVNAVSGVATLSSIVTGFINFKPFGGMLMIIIVIGLVFAIYFLTKTKIYETAP 791
QY 781 IKMIYPEIDKLKEREKSEIPISEELERIVLAMHIHQONSHMETKTRKPKDSILTRA 840
DB 792 IKMIYPEIDKLKEREKSEIPISEELERIVLAMHIHQONSHMETKTRKPKDSILTRA 851
QY 841 QNMLKRKSGYSLKNAESVEMLNTL 865
DB 852 QNMLKRKSGYSLKNAESVEMLNTL 876

RESULT 4
Q8B423
ID Q8B423 PRELIMINARY; PRT; 876 AA.
AC Q8B423;
DT 01-MAR-2003 (TReMBLrel. 23, Created)
DT 01-MAR-2003 (TReMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TReMBLrel. 23, Last annotation update)
DE Glycoprotein B.
OS Porcine lymphotropic herpesvirus 2.
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Gammaherpesvirinae.
OX NCBI_TaxID=91741;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=546;
RA Chmielewicz B., Goltz M., Franz T., Bauer C., Brema S., Ellerbrok H.,
RA Beckmann S., Rziha H.-J., Lahrmann K.-H., Romero C., Ehlers B.;
RT "A novel porcine gammaherpesvirus.";
RL Virology 0:0-0(2003).
DR EMBL; AY170314; AA012281.1; -.
SQ SEQUENCE 876 AA; 100771 MW; 3B238639221474C7 CRC64;

Query Match 93.8%; Score 4265; DB 12; Length 876;
Best Local Similarity 93.3%; Pred. No. 6e-260;
Matches 807; Conservative 28; Mismatches 30; Indels 0; Gaps 0;

QY 1 MAGSLKRGSLALWLYQVALYSLSIATGTGVTSPNTATWSTESPLTGHYTHDSSHGE 60
DB 12 MAGSLKRLALALWCQFQVALYSLKAEKATSANTQAWSTEALPIGQYGYDSSHGE 71
QY 61 RGNENRDSSEONKNIYGSPTFPYRVCSAGVGDFRFQTDHVCPCDASDMVHSEGIILLI 120
DB 72 RATSENREDEEHKNIYGSPTFPYRVCSAGVGDFRFQTDHVCPCDASDMVHSEGIILLI 131
QY 121 YKQNIIPFMRVRKRVKVTSTVYNGIYSDSIINQHTFYKSIPEWTEKMDTIYQCFNS 180
DB 132 YKQNIIPFMRVRKRVKVTSTVYNGIYSDSIINQHTFYKSIPEWTEKMDTIYQCFNS 191
QY 181 LRLNTGNNLLTYVDRDDINMTVFLQPDVGTDPVKRYGSOPELYLEPGWFGSYRRRTTV 240
DB 192 LRLNTGNNLLTYVDRDDINMTVFLQPDVGTDPVKRYGSOPELYLEPGWFGSYRRRTTV 251
```



```

OX NCBI_TaxID=10398;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=2148530; PubMed=11602790;
RA Dunowska M., Letchworth G.J., Collins J.K., DeMartini J.C.;
RT "Ovine herpesvirus-2 glycoprotein B sequences from tissues of ruminant
RT malignant catarrhal fever cases and healthy sheep are highly
RT conserved.";
RL J. Gen. Virol. 82:2785-2790(2001).
DR EMBL; AF385439; AAL24469.1; -.
DR InterPro; IPR000234; Glycoprot.B.
DR Pfam; PF00606; Glycoprotein_B; 1.
DR ProDom; PD000693; Glycoprot.B; 1.
SQ SEQUENCE 863 AA; 97529 MW; F7A8CCD71BDC32D9 CRC64;

Query Match          55.3%; Score 2513; DB 12; Length 863;
Best Local Similarity 55.7%; Pred. No. 1e-149;
Matches 476; Conservative 149; Mismatches 205; Indels 24; Gaps 6;

QY 21 ALYSLISAETGVTSPPTATWSTESPLTGHYTHDSSHGGERGNNENRDSEQKNIIYQSP 80
DB 15 ALLCLAACLGQTPPTPTT---EDDVIVPGHTVSPDLVKQQLGSLNEEBGSGINRGPM-DP 70
QY 81 STFPYRVCSAGVGDFRQTHVCPDASDMVHSEGLIIYKQNIIPMFVRVRYKRVVT 140
DB 71 SAFFPRVCSASNIGDIFRFQTHSCPTNKDKEHNEGILLIFKENIVPVYFKVRYKI 130
QY 141 TSTVYNGIYSDITNOHTFYKSIPEWETEKMDTIYOCNSLRNTGGLLTYYVDRDDINM 200
DB 131 TSTVYNGIYSDAITNOHESKSVPHYBARRMDTIYQCVNSLSTVGGNLLAYTDNDGYNL 190
QY 201 TVFLQPDVGVTPDKRYGSOPELYLEPGWFMGYSRRRTTVNCELMDMFARSPDPFVFT 260
DB 191 TVDLQPDGLSLSVRRYNSQPEIHAEPGLGGLYRRRTTVNCEVTDTEARSVPFRYFVT 250
QY 261 ATGDTVEMSPFWSGEGDDHKNKHEKPMFVSVNNYKVVDYQNRGTVPLGKTRIFLDREY 320
DB 251 NVGDTIEMSPFWSGGANETEPNKEPQRTVSVLRDYLVDYKDRGSRQPQHTRIFDKEDY 310
QY 321 TLSWEKHLKNNMYSYPLTLKAFYNGIQTEHSGSVHYFVANDITASFTTSKEDMKENFTYH 380
DB 311 TLSWAQQLKNNISYCRWAHWKSFHNAIKTEHNSYHFVANDITASFTTPTNTAEQDVTKTHT 370
QY 381 CLNEEIKAEIKKAKVKNSTHSGYKGLKYFKTDGGLYLVWQPLIQNLDAKNKL----- 435
DB 371 CLNLSIESEMTSRLEKYNHGTHTVNGSRQYLYLTNGGLLVWQPLVQQLLNAQDLLEAVAS 430
QY 436 -NNETVRRSRROAESRRAVSSILID-----DDVYTAESALLLTQIQFAYDMLRSQINTVLE 484
DB 431 KHNVTARSARRRQRRRAVSSILID-----DDVYTAESALLLTQIQFAYDMLRSQINTVLE 484
QY 495 DLKAWCREQHRRAALVNNELSKINPTSVMSMIYNRPVSAKRGDVISVNCIIVDQTSVS 554
DB 485 ELSRAWCREQHRRAALVNNELSKINPTSVMSMIYNRPVSAKRGDVISVNCIIVDQTSVS 544
QY 555 LHKSRLLSASD-EKCFSPPTVTFKPMNDSTYKQGLGVNNEILLTTTLETQCOENTFY 613
DB 545 LHRNMEVPCDRHAECYSPPVTFKFINDSHYLYKQGLGVNNEILLTTTLETQCOENTFY 604
QY 614 FOAKTDMYTYKNEYHLKTVPLSSITLDTFIALNFTLLENVDFKVIETYRDEKRLSNVF 673
DB 605 FQGGNNMYKYNRYHVKVTPISAVATLDTFIVNLTLVENIDFQVIELYSREKRMSTVF 664
QY 674 DIETPREYNYAQRVSGRLKOLLDLSTNRQFVAFGSLMDDLGAVGTVNAVSVAT 733
DB 665 DIETPREYNYTQRVTGLRSLDIATNRQFVAFGSLMDDLGAVGTVNAVSVAT 724
QY 734 LFSSIVTGFINFKIPFGGLMILIVVIGLFAIYELTKTKTIYETAPIKMIYPTDKLKE 793
DB 725 LFSSIVTGFINFKIPFGGLMILIVVIGLFAIYELTKTKTIYETAPIKMIYPTDKLKE 784
QY 794 RECKSEIAPISBEELERIVLAMHIHQONSHMETKTKDP-----KDSILTRAQNMRLK 846
DB 794 RECKSEIAPISBEELERIVLAMHIHQONSHMETKTKDP-----KDSILTRAQNMRLK 846
```

```

DB 785 QRQEMNVEPISKHELDRLIMLAMDHYHQTKQDKPDEKEGPESGGSANKANWLNAKVNLR 844
QY 847 RSGYSLNKNKAEVSVE 860
DB 845 RAGYQPLKRSNSTE 858

RESULT 7
QY1A90 PRELIMINARY; PRT; 863 AA.
AC QY1A90;
DT 01-DEC-2001 (Tremblrel. 19, Created)
DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)
DT 01-MAR-2002 (Tremblrel. 20, Last annotation update)
DE Glycoprotein B.
GN GB.
OS Ovine herpesvirus 2.
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Gammaherpesvirinae.
OX NCBI_TaxID=10398;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=2148530; PubMed=11602790;
RA Dunowska M., Letchworth G.J., Collins J.K., DeMartini J.C.;
RT "Ovine herpesvirus-2 glycoprotein B sequences from tissues of ruminant
RT malignant catarrhal fever cases and healthy sheep are highly
RT conserved.";
RL J. Gen. Virol. 82:2785-2790(2001).
DR EMBL; AF385441; AAL24471.1; -.
DR InterPro; IPR000234; Glycoprot.B.
DR Pfam; PF00606; Glycoprotein_B; 1.
DR ProDom; PD000693; Glycoprot.B; 1.
SQ SEQUENCE 863 AA; 97511 MW; 263EEB8B7870C20B CRC64;

Query Match          55.1%; Score 2504; DB 12; Length 863;
Best Local Similarity 55.5%; Pred. No. 3.9e-149;
Matches 474; Conservative 149; Mismatches 207; Indels 24; Gaps 6;

QY 21 ALYSLISAETGVTSPPTATWSTESPLTGHYTHDSSHGGERGNNENRDSEQKNIIYQSP 80
DB 15 ALLCLAACLGQTPPTPTT---EDDVIVPGHTVSPDLVKQQLGSLNEEBGSGINRGPM-DP 70
QY 81 STFPYRVCSAGVGDFRQTHVCPDASDMVHSEGLIIYKQNIIPMFVRVRYKRVVT 140
DB 71 SAFFPRVCSASNIGDIFRFQTHSCPTNKDKEHNEGILLIFKENIVPVYFKVRYKI 130
QY 141 TSTVYNGIYSDITNOHTFYKSIPEWETEKMDTIYQCVNSLRNTGGLLTYYVDRDDINM 200
DB 131 TSTVYNGIYSDAITNOHESKSVPHYBARRMDTIYQCVNSLSTVGGNLLAYTDNDGYNL 190
QY 201 TVFLQPDVGVTPDKRYGSOPELYLEPGWFMGYSRRRTTVNCELMDMFARSPDPFVFT 260
DB 191 TVDLQPDGLSLSVRRYNSQPEIHAEPGLGGLYRRRTTVNCEVTDTEARSVPFRYFVT 250
QY 261 ATGDTVEMSPFWSGEGDDHKNKHEKPMFVSVNNYKVVDYQNRGTVPLGKTRIFLDREY 320
DB 251 NVGDTIEMSPFWSGGANETEPNKEPQRTVSVLRDYLVDYKDRGSRQPQHTRIFDKEDY 310
QY 321 TLSWEKHLKNNMYSYPLTLKAFYNGIQTEHSGSVHYFVANDITASFTTSKEDMKENFTYH 380
DB 311 TLSWAQQLKNNISYCRWAHWKSFHNAIKTEHNSYHFVANDITASFTTPTNTAEQDVTKTHT 370
QY 381 CLNEEIKAEIKKAKVKNSTHSGYKGLKYFKTDGGLYLVWQPLIQNLDAKNKL----- 435
DB 371 CLNLSIESEMTSRLEKYNHGTHTVNGSRQYLYLTNGGLLVWQPLVQQLLNAQDLLEAVAS 430
QY 436 -NNETVRRSRROAESRRAVSSILID-----DDVYTAESALLLTQIQFAYDMLRSQINTVLE 484
DB 431 KHNVTARSARRRQRRRAVSSILID-----DDVYTAESALLLTQIQFAYDMLRSQINTVLE 484
QY 495 DLKAWCREQHRRAALVNNELSKINPTSVMSMIYNRPVSAKRGDVISVNCIIVDQTSVS 554
DB 485 ELSRAWCREQHRRAALVNNELSKINPTSVMSMIYNRPVSAKRGDVISVNCIIVDQTSVS 544
```

```

QY 555 LKSLRLLSASD-EKCFSPRPVTFKFMNDSTIYKGLQGVNNEILLTTTILETCOENTYY 613
D 545 LHRNRPVGRDHAHECVSRPPTFKFINDSHLYKGLQGVNNEILLTTTILEVCHENTHY 604
QY 614 FOAKTDMYIYKXNVEHLKTVPLUSSITTLDTFALNFTLLENVDFKVIELYTRDEKRLSNVF 673
D 605 FQGGNNMYFYKNYRHVKTIPVSAVATLDTFVLNLTIVENIDFQVIELYSREKRMSTVF 664
QY 674 DIETMREYNYAORVSLGRKDLDLSTNRNOFVDAFGSLMDDLGAQGVTVNAVSVAT 733
D 665 DIETMREYNYTORVGLRDLSDIATNRNQFVDAFGTLMDDLGVVGVKTVNAVSSLAT 724
QY 734 LFSSIVTGFINFKNPPFGGMLMIIVTGVLFVFAIYFLTKTKIYETAPIKMIYBEIDKLE 793
D 725 LFSSIVTGFINFKNPPFGGMLIFGLLAUVIAVILLRRAASFAANPVQMIYDIOQITK 784
QY 794 REGKSEIAPISBEELERIVLAMIHQONSHMETKTRKDP-----KDSILTRAQNLRLK 846
D 785 QRQEMNVEPIKSHELDRIMLAMHDYHQTQKDPDEKEGPESGGSANKANLNKAKNVLRR 844
QY 847 RSGVSNLKNAESVE 860
D 845 RAGYQPLKRSDEST 858

RESULT 8
Q91A89.
ID Q91A89 PRELIMINARY; PRT; 863 AA.
AC Q91A89;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DE 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Glycoprotein B.
GN GB.
OS Ovine herpesvirus 2.
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Gammaherpesvirinae.
OX NCBI_TaxID=10398;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21488530; PubMed=11602790;
RA Dunowska M., Letchworth G.J., Collins J.K., DeMartini J.C.;
RT "Ovine herpesvirus-2 glycoprotein B sequences from tissues of ruminant
RT malignant catarrhal fever cases and healthy sheep are highly
RT conserved.";
RL J. Gen. Virol. 82:2785-2790 (2001).
DR EMBL; AF385442; AAL24472.1; -.
DR InterPro; IPR000234; Glycoprot B.
DR Pfam; PF00606; Glycoprotein B; 1.
DR ProDom; PD000693; Glycoprot_B; 1.
SQ SEQUENCE 863 AA; 97534 MW; B2F63AB88876A777 CRC64;

Query Match 54.8%; Score 2493.5; DB 12; Length 863;
Best Local Similarity 55.4%; Pred. No. 1.8e-148;
Matches 475; Conservative 149; Mismatches 213; Indels 21; Gaps 5;

QY 17 LYQVALYSLSTAETAGVTSPPNTATWSTESPLTGHYGHDSHGEGNENRDSSEQKNVI 76
D 8 LSPAAITLLCLAAACGQTPTPTPTDDVIVPGNTVSPDVLKQOLGSGLNEEGSINRGP 67
QY 77 YGSPSTFPYRVCASGVGDVRFOTDHCVPDASDMVHSEGILLITYKONIIPFMFRVKYR 136
D 68 M-DPSAFPYRVCASASNGIDIFRQTHSCPNTKDKEHNEGILLIFKENIYVYFVKYR 126
QY 137 KVVTSTVYNGISDSTNQHFTFKYSIEPMEKMDTIYQCFNSRLNTCGNLLTYVD RD 196
D 127 KIVTSTVYNGISDAITNQHFSKSPHYEARMDTIYQCVNSLSLTGCGNLLAYTDND 186
QY 197 DINMTVFLQPDGVTPDKVYSGOPELYLPGHFWGSYRRRTTVNCELMDMFARPPPD 256
D 187 GYNLTVDLPQMDGLNSVRRNSQOPEIHAEPFGLLGGYRRRTTVNCEVDTTEARSVPFPR 246

```

```

QY 257 FFVTATGDTIVMSPFWSGEDDHENKHEKPFVSVINNVKVDYQNRGTVPGLKTRIFLD 316
D 247 YFVTNVGDITIEKSPFWSGGANETEPNKEPRRTVSVLRDYLTVDYKGRSPQPHTRIFID 306
QY 317 REBYTILSWKHLKMSYCPILTLWKAFYNGIQTEHSGSYHFVANDITASFTTSKEDMKFBN 376
D 307 KEDYTLSSWAQQLKNISYCRWAHWSFHNAIKTEHNSYHFVANDITASFTTPTEAQDVT 366
QY 377 TTVHCLNEEBIKABIEKKYAKVNSTHSHYKGLYFKTDGGLYLVMOPLIQNRLLDAKNKL- 435
D 367 KTHCLNSTIESBMTSRLEKVGTHVNGSRQYLLTNGGLLVQPLVQOQKLLNAQDLLE 426
QY 436 -----NNETYSRRSRQAESTTDPMMEMTGNAGGYSSENSITVAQOVAYDNLRIIN 490
D 427 AVASKENVTRESARSRRQRAVSSILID-----DDVYTAESALLTQIQFAYDMLRSQIN 480
QY 491 NILEDLSKAWCRQHRALVWNLSEKINPTSVMSIMYRNRPVSAKRIGDVISVSNICIVDQ 550
D 481 TVLEELSRAWCRQHRASLWNLSEKINPTSVMSIYGRPVSAKRIGDVISVSHCVVDQ 540
QY 551 TSVSLHKSRLLSASD-EKCFSPRPVTFKFMNDSTIYKGLQGVNNEILLTTTILETCOEN 609
D 541 QSVSLHRSRMVPGRDHAHECVSRPPTFKFINDSHLYKGLQGVNNEILLTTTILEVCHEN 600
QY 610 TEYFQAKTDMYIYKXNVEHLKTVPLSSITTLDTFFIALNFTLLENVDFKVIELYTRDEKRL 669
D 601 TEHYFOGNNMYFYKNYRHVKTIPVSAVATLDTFVLNLTIVENIDFQVIELYSREKRM 660
QY 670 SNVFDIETMREYNYAORVSLGRKDLDLSTNRNOFVDAFGSLMDDLGAQGVTVNAVS 729
D 661 STVFDIETMREYNYTORVGLRDLSDIATNRNQFVDAFGTLMDDLGVVGVKTVNAVS 720
QY 730 GVATLFSSIVTGFINFKNPPFGGMLMIIVTGVLFVFAIYFLTKTKIYETAPIKMIYBEID 789
D 721 SLATLFSSIVTGFINFKNPPFGGMLIFGLLAUVIAVILLRRAASFAANPVQMIYDIO 780
QY 790 KLKEREKSEIAPISBEELERIVLAMIHQONSHMETKTRKDP-----KDSILTRAQNL 842
D 781 QITKQEQMNVEPIKSHELDRIMLAMHDYHQTQKDPDEKEGPESGGSANKANLNKAKN 840
QY 843 MLRKSGVSNLKNAESVE 860
D 841 VLRRRAGYQPLKRSDEST 858

RESULT 9
Q91A91
ID Q91A91 PRELIMINARY; PRT; 863 AA.
AC Q91A91;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Glycoprotein B.
GN GB.
OS Ovine herpesvirus 2.
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Gammaherpesvirinae.
OX NCBI_TaxID=10398;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21488530; PubMed=11602790;
RA Dunowska M., Letchworth G.J., Collins J.K., DeMartini J.C.;
RT "Ovine herpesvirus-2 glycoprotein B sequences from tissues of ruminant
RT malignant catarrhal fever cases and healthy sheep are highly
RT conserved.";
RL J. Gen. Virol. 82:2785-2790 (2001).
DR EMBL; AF385440; AAL24470.1; -.
DR InterPro; IPR000234; Glycoprot B.
DR Pfam; PF00606; Glycoprotein B; 1.
DR ProDom; PD000693; Glycoprot_B; 1.
SQ SEQUENCE 863 AA; 97512 MW; 9464D7BD6DB863C1 CRC64;

Query Match 54.8%; Score 2491; DB 12; Length 863;

```

Best Local Similarity 55.6%; Pred. No. 2.5e-148; Matches 473; Conservative 145; Mismatches 199; Indels 34; Gaps 7;	
QY	24 SLISAETGVTSPNTATWSTESPLTGHYCTHDSHGERGNNENRDSEONKNIYGSPTF 83
Db	28 TLPTTDEDDVIVGNV-----SP-----DVLKQOLGSGUNESEGSINRGM-DPSAF 73
QY	84 PYRVCASAGVDFRQTDHVCPCDASDMVHSEGIILLIYKQNIIPFMRVRKYRKVVTTST 143
Db	74 PPRVCASASIGDIFRQTHSCPTNKDKEHNEGILLIPKENIVPVYFKVRKYRKIVTTST 133
QY	144 VYNGIYSDITNOHFYKSIPEWETEKMDTIYOCFNSLRNTGCGNLLTYVDRDDINMTVF 203
Db	134 VYNGIYSDAITNOHFESKVPYHEARRMDTIYQCNSLSLTVGGNLLAYTDNDGYNLTVD 193
QY	204 LQPVGVTPDVKRYGSOPELYLEPGWFGSYRRRTTNCCLMDMFARSPDPDFVTATG 263
Db	194 LQPMGLNSVRVYNSQPEIHAEPOWLLGGYRRRTTNCCEVTDTEARSVPPRYFVTNG 253
QY	264 DTVEMSPFWSGDDHDKHKKHKEPWFVSVINNVKVDYQNRGTVPGLKTRIFLDRBEYTLS 323
Db	254 DTIEMSPFWSGCGANETEPNKEPRRTVSLRDYTLVDYKDRGSRPOPHTRIFIDKEDYTLS 313
QY	324 WEKHLKMSYCPLLTWKAPYNGIQTEHSGSYHFVANDITASFTTSKEDMKENFTTYHCLN 383
Db	314 MAQOLKNISYCRWAHWSFHNAIKTEHENSYHFVANDITASFTTPEAQQDVTKTHTCLN 373
QY	384 EBIKAEIEKKYAKVNSTHSGYDGLYFKTDGGLYLWVQPLIQNLRLDLDKKNL-----NN 437
Db	374 SLIESEMTSRLEKVGTHVNGSRQYLLTNGGLLLWVQPLWQOOLNLODLEAVASKHN 433
QY	438 ETSYRSRRQABSTTDPMMEMTCGAGGEYSSENSITVAQVQYAYDNLRIINNILEDLS 497
Db	434 VTKPARGRRQRAVSSILID-----DDVYTAESALLTQIQAFVDMLSQINTVLEELS 487
QY	498 KAWCREQHRASLWNLNLSKINPTSVMSIYNRPVSASAKRIGDVISVNCIYVDDQTSVSLHK 557
Db	488 RAWCREQHRASLWNLNLSKINPTSVMSIYGRPVSAKRIGDVISVSHCVVDDQQSVPLHR 547
QY	558 SLRLASD-EKCFSPRPVTFKFMNDSTIYKQOLGVNNEILLITTYLETQENTYFYQQA 616
Db	548 NMRVPGRHAHECSPRPVTFKFMNDSTIYKQOLGVNNEILLITTYLETQENTYFYQQA 607
QY	617 KTDVIYKNEYHLKTVPLASSITTLDTFIALNFTLLENVDFKVIELYTRDEKRLSNVFDIE 676
Db	608 GNNMYFYKNYRHVKTIPIPSAVATLDTFVLNLTVENIDFQVIELYSREKRMSTVFDIE 667
QY	677 TMFREYNYAQRVSGRLKDLDLSTNRNQFVDAFGSLMDLGAQGQTVVNAVSGVATLFS 736
Db	668 TMFREYNYTQRTVGLRDLSDIATNRNQFVDAFGTLMDDLGVGKTVVNAVSSLATLFS 727
QY	737 SVTGGFINFKNPPFGMLMIIVIGVFAIYFLTKTIYETAPIKMIYPEIDKUKEREG 796
Db	728 SVTGLINFINFKNPPFGMLIFGLAAVIAVILLRRRAAFANPQMIYPDQIQQITKQK 787
QY	797 KSEIAPISEEELERIVLAMIHIQNSHMETKRPD-----KDSILTRAQNLMLKRESG 849
Db	788 ENNVFPISEKHELDRLIMLAMDHDTQKDPDEKEGSPGSGSANKANLWLNKAKNVLRRAG 847
QY	850 YSNLKNAESVE 860
Db	848 YQPLKRSOSTE 858
RESULT 10	
O36362	
ID	O36362
AC	O36362;
DT	01-JAN-1998 (TReMBLrel. 05, Created)
DT	01-JAN-1998 (TReMBLrel. 05, Last sequence update)
DT	01-JUN-2001 (TReMBLrel. 17, Last annotation update)
DE	Glycoprotein B.
OS	Alcelaphine herpesvirus 1 (wildbeest herpesvirus).
Viruses; dsDNA viruses, no RNA stage; Herpesviridae; Gammaherpesvirinae; Rhadinovirus. NCBI_TaxID=35252; [1]_TaxID=35252; SEQUENCE FROM N.A. STRAIN=CS00; MEDLINE=97404659; PubMed=9261371; R. Enser A., Pflanz R., Fleckenstein B.; "Primary structure of the alcelaphine herpesvirus 1 genome."; J. Virol. 71:6517-6525 (1997). EMBL; AF005370; AAC58059.1; -- InterPro; IPR000234; Glycoprot. B. Pfam; PF00606; Glycoprotein B; 1. ProDom; PD00693; Glycoprot B; 1. SEQUENCE 854 AA; 97059 MW; 480D7CF352D5B59C CRC64;	
Query Match 54.8%; Score 2489.5; DB 12; Length 854; Best Local Similarity 57.0%; Pred. No. 3.1e-148; Matches 471; Conservative 135; Mismatches 173; Indels 47; Gaps 8;	
QY	64 NENRDSBEON---KNYGSPTPEYRVCASAGVDFRQTDHVCPCDASDMVHSEGIILLI 120
Db	42 SDNIQRONNATKAGIHDPSAFPFRVCASASIGDIFRQTHSCPTNKDKEHNEGILLI 101
QY	121 YKQNIIPFMRVRKYRKVVTTSTVYNGIYSDITNOHFYKSIPEWETEKMDTIYOCFNS 180
Db	102 FKENIVPVYFKVRKYRKIVTTSTIYNGIYADAVTNQHVFSKSVPIYETRRMDTIYQCNS 161
QY	181 LRLNTGNNLLTYVDRDDINMTVFLQPVGVTPDVKRYGSOPELYLEPGWFGSYRRRTTV 240
Db	162 LDVTVGGNLLTYVDRDDINMTVFLQPVGVTPDVKRYGSOPELYLEPGWFGSYRRRTTV 221
QY	241 NCELMDMFARSPDPDFVTATGTVEMSPFWSKAMNETEFGSDPDT-----SGDDHKNKHEKHPWFSV 291
Db	222 NCEVTDARAVPFRFYFITNIGDTIEMSPFWSKAMNETEFGSDPDT-----LTV 272
QY	292 INNVKVDYQNRGTVPGLKTRIFLDRBEYTLSEKHLKMSYCPLLTWKAPYNGIOTSHS 351
Db	273 AKDVRVDYKFRGTQPOGHTRFVDKBEYTLWQAQFENISYCRWAHWSKFDNAIKTEHG 332
QY	352 GSYHFVANDITASFTTSKEDMKENFTTYHCLNEBIKAEIEKKYAKVNSTHSGYDGLYFK 411
Db	333 KSLHFVANDITASFTTNTQTRVLGKHVCLNNTIESELKSLRLAKVNDTHSPNGTAQYVL 392
QY	412 TDGGLYLWVQPLIQNLRLDLDKKNL-----NETYSRSRRQABSTTDPMMEMTCG 462
Db	393 TNGGLYLWVQPLWQOOLNLODLEAVASKHN-----SG 444
QY	463 AGGEYSSENSITVAQVQYAYDNLRIINNILEDLSKAWCREHRAALWNLNLSKINPTSV 522
Db	445 IDVYTAESTILLTQIQAFVDMTLRAQINNVLEELSRACREHRAALWNLNLSKINPTSV 504
QY	523 MSMTYNRPVSASAKRIGDVISVNCIYVDDQTSVSLHKSRLLSA-SDEKCFSPRPVTFKPMN 581
Db	505 MSSYGRPVSAKRIGDVISVSHCVVDDQTSVSLHSMRVPGDKTHECYSPRPVTFKFIN 564
QY	582 DSTIYKQOLGVNNEILLITTYLETQENTYFYQQAQTDVIYKNEYHLKTVPLSSITLTD 641
Db	565 DSHLYKQOLGVNNEILLITTYLETQENTYFYQQAQTDVIYKNEYHLKTVPLSSITLTD 624
QY	642 TFIALNFTLLENVDFKVIELYTRDEKRLSNVFDIETMFREYNYAQRVSGRLKDLDLST 701
Db	625 TFMVNLTLVENIDFQVIELYSREKRMSTAFDIETMFREYNYTQRTVGLRDLDTLAT 684
QY	702 NRNQFVDAFGSLMDLGAQGQTVVNAVSGVATLFSSTVITGFINFKNPPFGMLMIIVIG 761
Db	685 NRNQFVDAFGSLMDLGAQGQTVVNAVSGVATLFSSTVITGFINFKNPPFGMLLFLGLIAA 744
QY	762 VLFPAIYFLTKTIYETAPIKMIYPEIDKUKERCKSEIAPISEEELERIVLAMIHIQNS 821
Db	745 VVITVILLNRKAKFAQNPQMIYPDKITTSQREELQVDPISKHELDRLIMLAMDHYSAS 804
QY	822 SHMETK-----TRKDPKDSILTRAQNLMLKRESGYSNLKNAESVE 860


```

QY 366 TTSDKEMKFNFTYHCLNEIEKAEIKYKAVNSTHSHKYGDLKFKTKDGLYLWQPLIQ 425
D 338 STPLEEENFNFTFCWNTQEEIKKLEKEVETHRPNTAKVTKTGNLYVWQPLIQ 397
QY 426 NLLDAKUKLNN-----FTYRRSRROAESTTDPMMWTGAGGEGYS-ENS 472
D 398 IDLDTHTAKLYNLTNATASPTSTPTSPRRRRDTS-----VSGGNGNDNSTKEES 450
QY 473 ITVAQVQAYDNLRIINILIEDLSKAWCREQHRALVNLNLSKINPTSVMSMIYRNPVS 532
D 451 VAASQVQFAYDNLRSINRVLGELSRACREQYRASLMWYELSKINPTSVMSAIYGRPVS 510
QY 533 AKRIGDVISVNCIIVDQTSVSLHKSRLLSASDEKCFSPRPVTEFKFVNDSTIYKQGLGV 592
D 511 AKLIGDVSVSDCISVDQKSVFVHKMK-VPCKEDLCYTRPVVGFKINGSELPAGQLGP 569
QY 593 NNEILLTTLTTCENTYFQAKTDMYIKNYEHLKTVPLSSITTLDTFIALNFTLLE 652
D 570 RNEIYVLSQVEVCQSHCEHYFQAGNOMYKYKDYVYVSTLNTDPTLHTMITLNLSLVE 629
QY 653 NVDFKVIELYTRDEKRLNVFDIETMFREYNYYAQRVSLRKDLAD-LSTNRNQFVDAFG 711
D 630 NIDFKVIELYSKTERLSNVFDIETMFREYNYYTQNLGRLKDLDSIDHGRDSTIQTIG 689
QY 712 SLMDLGAQVQTVNAVSGVATLFSSTVTFINFKNPFPGMLIIVVI-GVLEFAYELT 770
D 690 DIMDLGTIGKVVNVASGVSLFGSIVSGVIFKFNPFPGMLLIVLIAGVVVYLFMT 749
QY 771 KTKIYETAPIMYIPEIDKUKEREGKSEIAPISEEELERIVLAMHIHQONSHMETKTRK 830
D 750 RRSIY-SAPIRMLYPGVERAAQEPGAH---PVSEDIQRNLMGMHQFQORAEAEARR 805
QY 831 DPK-----DSILTRAQNMRLKR---SGYSNLK 854
D 806 EEEVKGKTLFEVIRDSATSVLRRRRGGGQYRLQ 840

RESULT 13
P88906
AC P88906; PRELIMINARY; PRT; 845 AA.
DT 01-MAY-1997 (TReMBLrel. 03, Created)
DT 01-MAY-1997 (TReMBLrel. 03, Last sequence update)
DT 01-JUN-2001 (TReMBLrel. 17, Last annotation update)
DE ORF 8.
OS Kaposi's sarcoma-associated herpesvirus (KSHV) (Human herpesvirus 8).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Gammaherpesvirinae; Rhadinovirus.
OX NCBI TaxID=37296;
RN [1]_
RN SEQUENCE FROM N.A.
RX MEDLINE=97121480; PubMed=8962146;
RA Russo J.J., Bohenzky R.A., Chen M.C., Chen J., Yan M., Maddalena D.,
RA Parry J.P., Peruzzi D., Edelman I.S., Chang Y., Moore P.S.;
RT "Nucleotide sequence of the Kaposi sarcoma-associated herpesvirus
RT (HHV8)".
RL Proc. Natl. Acad. Sci. U.S.A. 93:14862-14867 (1996).
RN [2]
RN SEQUENCE FROM N.A.
RX MEDLINE=97094384; PubMed=8939871;
RA Moore P.S., Bashoff C., Weiss R.A., Chang Y.;
RT "Molecular mimicry of human cytokine and cytokine response pathway
RT genes by KSHV".
RL Science 274:1739-1744 (1996).
RN [3]
RN SEQUENCE FROM N.A.
RX MEDLINE=97296220; PubMed=9151804;
RA Neipel F., Albrecht J.C., Fleckenstein B.;
RT "Cell-homologous genes in the Kaposi's sarcoma-associated rhadinovirus
RT human herpesvirus 8: determinants of its pathogenicity?";
RL J. Virol. 71:4187-4192 (1997).
RN [4]

```

```

RP SEQUENCE FROM N.A.
RT Pertel P.E., Spear P.G., Longnecker R.;
RT "Human herpesvirus-8 glycoprotein B interacts with Epstein-Barr virus
RT (EBV) glycoprotein 110 but fails to complement the infectivity of EBV
RT mutants.";
RL Virology 0:0-0 (1998).
DR EMBL; U75698; AAC57085.1; -.
DR EMBL; U93872; AAB62592.1; -.
DR EMBL; AF092928; AAC83368.1; -.
DR InterPro; IP000234; Glycoprot_B.
DR Pfam; PF00606; Glycoprotein_B; 1.
DR ProDom; PD00693; Glycoprot_B; 1.
SQ SEQUENCE 845 AA; 93984 MW; C4A5F2FF29B1E41 CRC64;

Query Match 42.6%; Score 1937; DB 12; Length 845;
Best Local Similarity 48.5%; Pred. No. 1.8e-113;
Matches 382; Conservative 135; Mismatches 237; Indels 34; Gaps 10;

QY 83 FPYRCSASGVGVDFRFQTDHVCPCDASDMVHSEGILLIYKQNIIPFMRVKRYKVVTTTS 142
D 63 YQFVCSASITGELFRFNLEQTCPTDKYHQEGILLVYKKNIVPHIFKVRRYKRIATSV 122
QY 143 TVYNGIYSDSITNOHTFYKSIPEWETEMOTIYOCFNSLRNLNTGNNLLTYVDRODINMTV 202
D 123 TVYRGLTESAITNKYELPRPVLYEISHMDSTYQCFSSMKYVNVGVENTFTDRDDVNTTV 182
QY 203 FLQPVGVGTPDVKRYGSOPELYLEPGFWGYSRRTTVCCELMDFARSNPPFPFVTAT 262
D 183 FLQPEVGUDNIQRYFQPIYABPGFWPPGPIYRVTTCVCEIVDMIARSAEPYNYFVTSL 242
QY 263 GDTVEMSPFWSGEDDHENKHEKMPF-VSVINNYKVVDYQNRGTVPGLKTRIFLDRBEYTT 321
D 243 GDTVEVSPFCVNESSCSTPSNKNGLSVQVVLNHTVTVSDRGTSPTPQNPFIIVETGAYT 302
QY 322 LSWEKHLKNMSYCPCLTLWAPYNGIQTEHSGSYHFVANDITASPTTSKEDMKEFTTYVHC 381
D 303 LSWASESKTTAVCPALWKTTPRSIQTTTHEDSFHFVANEITATPTAPLTPVANFTDYSC 362
QY 382 LNEEIKAEIEKKYAKVNSTHSHKYGDLKFKTKDGLYLWQPLIQNLRLDANKLNNET-- 439
D 363 LTSDINTLNLASKAKLASTHVPNGTVQYFHTTGLYLWQPMASINLTHAOGDSGNPTSS 422
QY 440 -----YSRRSRROAESTTDPMMWTGAGGEGYSSENSITVAQVQYADNLRI 488
D 423 PPSASPMWTSASRRKRSASTA-----NAGGGGSTDN-LSYTLQFADYDKLRDG 471
QY 489 INNILEDLSKAWCREQHRALVNLNLSKINPTSVMSMIYRNPVSAGRIDGVISVNCIIV 548
D 472 INQVLEELSRACREQVRDNLWYELSKINPTSVMTAIYGRPVSAKFGVAISVTECINV 531
QY 549 DOTSVSLHKSRLLSASDEKCFSPRPVTEFKFEMNDSTIYKQGLVNNELLTTLTTCQE 608
D 532 DQSSVNIHKSRL--TNSKDVCYARPLVTFKLNSSNLTGOLGARNIEILTNNOVETCKD 589
QY 609 NTEYFQAKTDMYIKNYEHLKTVPLSSITTLDTFIALNFTLNVDPFKVIELYTRDEKR 668
D 590 TCEHYFTRNETLVYKQYAYLRTINTDITLNTFIALNLSFQINIDPKAIELYSSAEKR 649
QY 669 L-SNVFDIETMFREYNYYAQRVSLRKDLAD-LSTNRNQFVDAFGSLMDLGAQVQTVN 726
D 650 LASSVFDLETMFREYNYYTHRLAGREDLONTIDMKERFVRDLSEIVADLGGIGKTVN 709
QY 727 AVSGVATLFSSTVTFINFKNPFPGMLIIVIGVLFAIYFLTKTKTIYETAPIKMIYP 786
D 710 VASSVWTLGSLVTGTFINFKHPLGGLMHIIVIAIILIIIFMLSRRTNTIAQAPVKMIYP 769
QY 787 EIDKLKEREGKSEIAPISEEELERIVLAMHIHQONSHMETKTRKDPKDSILTTRAQNMRLK 846
D 770 DVDR---RAPSPGAP-TREEIKNILLGMHQLQOEERQKADDLKSTPSVQRTANGLRQ 825
QY 847 R-SGYSNLK 853
D 826 RLRCYKPL 833

```



```

QY 293 NNYKVVYQNRCTVPLGKTRIFLDREBYTSLWEKHLKMSYCPLTLWKAFYNGIQTEHSG 352
Db 270 TNYTVDVFATR--APTETRVFADSGEYTVSMKAEDPKSAVCALTLWKTFPRAIQTTHEA 327
QY 353 SYHFVANDITASFTTSKEDMKFNNTYHCLNEEIKAEIEKKYAKVNSTHYSKYGDUKYFKT 412
Db 328 SYHFVANDVATFTSPLSEVANFTYSCLDDEVIOKTLNDTIKKLSDTHVNGSAQYKYKT 387
QY 413 DGGLYLVWQPLTONRLDANKLNNETYS-----RRSRQAESTTDPMMEMTGN 461
Db 388 EGGFLWQPLTSLVDEMRLNGTTTAPPATTSTVSRVRSVNTNEQATD----- 439
QY 462 GAGGEYSSENSITVAQVOYAYDNLRIRINNILEDSKAWCREQHRALVWNELSKINPTS 521
Db 440 -----NLAAPOLOFAYDKLRASINKVLEELSRAWCREQVRDITYMYELSKINPTS 489
QY 522 VMSMIYNRPVSARKRIGDIVISVNCIIVDQTSVLSLHLSKLSASDEKCFSPRPVTFKPMN 581
Db 490 VMTAIYGRPVSAKFVGDALSVTDCVAVDQASVSIHKSLR--TSTPGMCYSRPPVTFRFLN 547
QY 582 DSTIYKQGLGVNNEILLTYYLETQCENTEYVFOAKTDMYIKNYEHLKTVPLSSITLTD 641
Db 548 STTLFKGQLGPRNEIILTNDQVEACKETCEHYFIASNTVYKQVYFVKKINTSEISLFG 607
QY 642 TFIALNFTLLENVDFKVIELYTRDEKRLS--NVFDIETMFREYNYAQRVSGLRKDL---L 697
Db 608 TFIALNLSFIENIDFRVIELYSRAEKLGSVFDIETMFREYNYTQRLAGLREDLDNTI 667
QY 698 DLSTNRNQFVDAGFGLMDLDGAVGOTVYVNAVSGVATLFSISVTGFINIKNPPFGMLMI 757
Db 668 DL--NRDLARDLSEIVADLGDVGRTVVNAVSSVITLFGSIVSGFINIKSPFGGMLMIL 725
QY 758 VVIGVLPFAIYFTKTKTIYETAPIKMIYPEIDKLEREGKSEIAPISEEELERIVLAMHI 817
Db 726 VIVAVLVFALNRRTNAIAQAPIRMYPIDDKMQPSGGK-----VDOEQIKNLAGMHQ 780
QY 818 HQONSHMETKTRDKPDKOSILTRAQNMRLKR--SGYSNLKNABSVE 860
Db 781 LQOEERRRLDEQORSAPSLFRASDGLKRRFRGYKPLENEAEQ 824

RESULT 16
Q9J2M4 PRELIMINARY; PRT; 829 AA.
ID AC Q9J2M4
DT 01-OCT-2000 (TReMBLrel. 15, Created)
DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)
DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
DE ORF08.
OS Macaca mulatta rhadinovirus 26-95.
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Gammaherpesvirinae; Rhadinovirus.
OX NCBI_TaxID=119193;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Macaca mulatta rhadinovirus isolate 26-95;
RX MEDLINE=20173730; PubMed=10708456;
RA Alexander L., Denekamp L., Knapp A., Auerbach M.R., Damanian B.,
RA Desrosiers R.C.;
RT "The primary sequence of rhesus monkey rhadinovirus isolate 26-95:
RT sequence similarities to Kaposi's sarcoma-associated herpesvirus and
RT rhesus monkey rhadinovirus isolate 17577.";
RL J. Virol. 74:3388-3398(2000).
DR EMBL; AF210726; AAF59985.1; -.
DR InterPro; IPR000234; Glycoprot_B.
DR Pfam; PF00606; Glycoprotein B; 1.
DR ProDom; PD000693; Glycoprot B; 1.
SQ SEQUENCE 829 AA; 93230 MW; 1459F972DB8B9EF4 CRC64;

```

Query Match 42.3%; Score 1925.5; DB 12; Length 829;
 Best Local Similarity 44.2%; Pred. No. 9.3e-113;
 Matches 391; Conservative 144; Mismatches 246; Indels 103; Gaps 15;

```

QY 11 VLALWLYQVALYSLSIAETCTVSPNTATWTSTESPLTGHYGTHTDSSHGSGNNENROSE 70
Db 10 LLRAVW--ITAIGTAVGE--NVTTPKGATTAKATFP----- 42
QY 71 EQNKNIYGSPT-----FPYRVCSASGVGVFRFQTDHVHCPDASDMVHSEGIL 118
Db 43 -----GPSTPTPPENPPRAEAFKFRVCASASATGELFRFNLEKTCFGETDKTHQEGIL 94
QY 119 LIYKONILPFMRVRKRYKVVTTSTVYNGIYSDSITNQHTFYKSIPEWETEMKDTIYOCF 178
Db 95 MVFKKNIVPHIFKRYRKYKVAITSVTYRGMVTETAVTGKQEVIRPVPOYEINHMDTTYOCF 154
QY 179 NSLRINTGGNLLTYVDRDDINMTVFLQPVDCVTPDVKRYGSOPELYLEPGHFWGMSYRRRT 238
Db 155 SSMVNVNGIVNTYTRDFTNQTVFLQPVDELTDNIQRYFSQPVLYITPFGHFGIYRVRT 214
QY 239 TVNCELMDMFARSNPPFFDFTVATGDTVEMSPFWSGE-----DDHENKMKHEKPMFVSVI 292
Db 215 TVNCEIVDMIARSAEPPSYFVTALGDTVEVSPPFCHNDSTCSVAEKTENGLG-----VRAL 269
QY 293 NNYKVVYQNRCTVPLGKTRIFLDREBYTSLWEKHLKMSYCPLTLWKAFYNGIQTEHSG 352
Db 270 TNYTVDVFATR--APTETRVFADSGEYTVSMKAEDPKSAVCALTLWKTFPRAIQTTHEA 327
QY 353 SYHFVANDITASFTTSKEDMKFNNTYHCLNEEIKAEIEKKYAKVNSTHYSKYGDUKYFKT 412
Db 328 SYHFVANDVATFTSPLSEVANFTYSCLDDEVIOKTLNDTIKKLSDTHVNGSAQYKYKT 387
QY 413 DGGLYLVWQPLTONRLDANKLNNETYS-----RRSRQAESTTDPMMEMTGN 461
Db 388 EGGFLWQPLTSLVDEMRLNGTTTAPPATTSTVSRVRSVNTNEQATD----- 439
QY 462 GAGGEYSSENSITVAQVOYAYDNLRIRINNILEDSKAWCREQHRALVWNELSKINPTS 521
Db 440 -----NLAAPOLOFAYDKLRASINKVLEELSRAWCREQVRDITYMYELSKINPTS 489
QY 522 VMSMIYNRPVSARKRIGDIVISVNCIIVDQTSVLSLHLSKLSASDEKCFSPRPVTFKPMN 581
Db 490 VMTAIYGRPVSAKFVGDALSVTDCVAVDQASVSIHKSLR--TSTPGMCYSRPPVTFRFLN 547
QY 582 DSTIYKQGLGVNNEILLTYYLETQCENTEYVFOAKTDMYIKNYEHLKTVPLSSITLTD 641
Db 548 STTLFKGQLGPRNEIILTNDQVEACKETCEHYFIASNTVYKQVYFVKKINTSEISLFG 607
QY 642 TFIALNFTLLENVDFKVIELYTRDEKRLS--NVFDIETMFREYNYAQRVSGLRKDL---L 697
Db 608 TFIALNLSFIENIDFRVIELYSRAEKLGSVFDIETMFREYNYTQRLAGLREDLDNTI 667
QY 698 DLSTNRNQFVDAGFGLMDLDGAVGOTVYVNAVSGVATLFSISVTGFINIKNPPFGMLMI 757
Db 668 DL--NRDLARDLSEIVADLGDVGRTVVNAVSSVITLFGSIVSGFINIKSPFGGMLMIL 725
QY 758 VVIGVLPFAIYFTKTKTIYETAPIKMIYPEIDKLEREGKSEIAPISEEELERIVLAMHI 817
Db 726 VIVAVLVFALNRRTNAIAQAPIRMYPIDDKMQPSGGK-----VDOEQIKNLAGMHQ 780
QY 818 HQONSHMETKTRDKPDKOSILTRAQNMRLKR--SGYSNLKNABSVE 860
Db 781 LQOEERRRLDEQORSAPSLFRASDGLKRRFRGYKPLENEAEQ 824

RESULT 17
Q9YTO5 PRELIMINARY; PRT; 807 AA.
ID AC Q9YTO5
DT 01-MAY-1999 (TReMBLrel. 10, Created)
DT 01-MAY-1999 (TReMBLrel. 10, Last sequence update)
DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
DE Glycoprotein B.
OS Arelaine herpesvirus 3.
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Gammaherpesvirinae; Rhadinovirus.

```



```

QY 604 ETCQENTYFQAKTDMYIKYNEHLKTVPLSSITLTFIALNFTLLENVDFKVIELYT 663
Db 577 EVCQHSCEHYFQGNHIYKYKNEYVSTWNLTDVPLTMTITLNLSLVENVDFQVQLVS 636
QY 664 RDEKLSNVDFIETMFREYNYAQRVGLRKDLDD-LSTNRNQFVDAGSLMDLGAQGO 722
Db 637 QKEKLSNVDFIETMFREYNYITQNLGRKDLDDSIHGRDSFTQLGLDVLQDLPVPGD 696
QY 723 TVNVASGVATLFSISITVGFINFKNPFGGMLMIIVIGLVFAIY-FLTKTKIYETAPI 781
Db 697 VIVNVASGVFSLFGSIVSGVISFLKNPLGAILITIALIVGGIIVLFIETRSVTYQ-API 755
QY 782 KMIYPEIDLKREGKESEIAPISEBELERIVLAMIHQNSHMETKTRDKPK-----D 834
Db 756 RMLYPEVD---RAPQNVQPIPEQVRSILLAMEHQFOQOQOQOQOQOQOQOQOQOQO 811
QY 835 SILTRAQNMRLKX---SGSYNLKNAES 858
Db 812 TIRESTSNILRRRRGGGYTRLRQDS 838

RESULT 19
Q93K7
ID Q93K7 PRELIMINARY; PRT; 870 AA.
AC Q93K7;
DT 01-JUN-2001 (TremBrel. 17, Created)
DT 01-JUN-2001 (TremBrel. 17, Last sequence update)
DT 01-DEC-2001 (TremBrel. 19, Last annotation update)
DE ORF4
OS Calitrichine herpesvirus 3.
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Gamaherpesvirinae; Lymphocryptovirus.
OX NCBI_TaxID=106331;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CJ0149;
RX MEDLINE=21107697; PubMed=11158621;
RA Cho Y., Ramey J., Rivallier P., Quink C., Garber R.L., Beier D.R.,
  Wang F.;
  "An Epstein-Barr-related herpesvirus from marmoset lymphomas.";
  Proc. Natl. Acad. Sci. U.S.A. 98:1224-1229(2001).
RL EMBL; AF319782; AAK38211.1; -.
DR Interprot; IPRO00234; Glycoprot B.
DR Pfam; PF00606; Glycoprotein B; 1.
DR ProDom; PD000693; Glycoprot B; 1.
SQ SEQUENCE 870 AA; 98634 MW; 6B7F0E0A7A0E8450 CRC64;

Query Match 40.5%; Score 1841; DB 12; Length 870;
Best Local Similarity 44.3%; Pred. No. 2.1e-107;
Matches 371; Conservative 164; Mismatches 233; Indels 70; Gaps 17;

QY 80 ESTFYRVCASGVGDVFRFQTDHVCPPAS-DMVHSEGILLIYKQNIIPFMRVRYKRV 138
Db 45 PYEFPFRVCEMNHGDIIFRATRISCPNFNLQENHTEGLIVFKDNIIPFNVRAYTKV 104
QY 139 VTTSTWNGIYSDSITNHTFYKSTEPWETEKMDTIYOCENSLRLNTGNTLLTYVDRDDI 198
Db 105 VTTVTWNGWYADAVTNAHEERYSPDYLENNMDTIYOCYNSVRMVKAGVERVYVDRDK 164
QY 199 NMTVFLQPDVGPDPVKRYGSOPELYLEPGFWGYSRRRTTVCNCLMDMFARSNPPDF 258
Db 165 NITVNLRPASGISTGVRRVNSQDLYDDPGWVWYRRTTVCNMTDMQAKSNSPFF 224
QY 259 VTATGDTVMSPFWGSDHDKMKWPFVSVINNYKVQYQNGRTVPLGKTRIFLDRE 318
Db 225 VTTLGQTVMSPFYNG--SNSETFNEPAGTFEIREKYIVTPENRGITIPKEKRAFLNKD 282
QY 319 EYTLSEKHLKMWSCPLTLKAFYNGIOTESHGSHYFVANDITASTFTSKEDMKFNTT 378
Db 283 TTYLSSKLDARS-SYCLPTMMKDYSAIRTDGSGSFHPTPOGTASFTV-----NTT 333
QY 379 -----YHCLNEEKAEIEKKYAKVNSTHSA-YGDLKYFKTDGGLVYLWQPLIQNRLD 430

```

```

Db 334 HEKLPDQFKCIEQQVKNITTEQFAKLSKNKYDRDAEIQYFRTAGCLLAWLPVTPKSLIT 393
QY 431 AKNNLNNETYS-----RESRROAES-----TTPDMNEMTGNG 462
Db 394 IRDIALNNTAQDGAARDQPDPTANSPTPOAPLRNRRRAGSDRNTNTITDVSATTEG 453
QY 463 AGGEYSSENSI---TVAQVQVAYDNLRIRINNILEDLSKAWCRQHRALVYNELSKINP 519
Db 454 TGDKDSILGTIDNPATAIQIQAIDSLRSQINRMGLDGLASAWCQORRQNMVLNLTKNP 513
QY 520 TSVMSIYNRPVSAKRIGDVISVNSCIVDQTSVLSLHSLRLSASDEKCSRPVTPKF 579
Db 514 TTVMTGVYGPAAKRIGDVISVRCVPVEQESVSLRSMRVL-GSETSCYTRPLVSFSF 572
QY 580 MNDSTIYKGLQGVNNEILLTTTLETQENTYFQAKTDMYIKYNEHLKTVPLSSITT 639
Db 573 PDSKTYEGQGVQNEILLTKMVENCQDTQCHTFQSGNEMHVRFDYQHFXTIPEVDVAT 632
QY 640 LDTFIALNFTLLENVDFKVIELYTRDEKRLSNVFDIETMFREYNYAQRVSGLRKDLDD- 698
Db 633 LOTFITLNTSFITNIDFQTLLEYSDRERRASNVFDIEGIFREYNFQTNIANLRDLDA 692
QY 699 LSTNRNQFVDAFGSIMDDDLGAVQTVNVASGVATLFSISITVGFINFKNPFGGMLMIIV 758
Db 693 ISNNRQFVDGLGELTDSLGAVGQAITNIVSSVGLFSLISGFSIFPKNPFQGVLLIAG 752
QY 759 VIGVLEAIYELTKTKIYETAPIKMIYPEIDLK-KERE-GKSEIAPISEBELERIVLAMI 816
Db 753 IVAVIFLVIFLVQRNRTIAQOPVQFLYPEIQSLTKNREIQGSNVAPISKQDLAIMLALY 812
QY 817 IH-QQNSHMETKTRKPKDSILTRA-----QNNMLKRSGYSN-----LKNAESVE 860*
Db 813 EQTKRTQAESKT-SDSTTSLPSRALEARNRLRKKPGRENRLKPLLTNIEDTE 869

RESULT 20
Q8UZD5
ID Q8UZD5 PRELIMINARY; PRT; 864 AA.
AC Q8UZD5;
DT 01-MAR-2002 (TremBrel. 20, Created)
DT 01-MAR-2002 (TremBrel. 20, Last sequence update)
DT 01-JUN-2002 (TremBrel. 21, Last annotation update)
DE BALF4.
OS Cercopithecine herpesvirus 15.
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Gamaherpesvirinae; Lymphocryptovirus.
OX NCBI_TaxID=104228;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=LCL8664;
RX MEDLINE=97048062; PubMed=8892903;
RA Franken M., Devergne O., Rosenzweig M., Annis B., Kieff E., Wang F.;
  "Comparative analysis identifies conserved tumor necrosis factor
  receptor-associated factor 3 binding sites in the human and simian
  Epstein-Barr virus oncogene LMP1.";
  J. Virol. 70:7819-7826(1996).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=LCL8664;
RX MEDLINE=99412410; PubMed=10482645;
RA Rivallier P., Quink C., Wang F.;
  "Strong selective pressure for evolution of an Epstein-Barr virus
  LMP2B homologue in the rhesus lymphocryptovirus.";
  J. Virol. 73:8867-8872(1999).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=LCL8664;
RX MEDLINE=20304984; PubMed=10846073;
RA Jiang H., Cho Y.G., Wang F.;
  "Structural, functional, and genetic comparisons of Epstein-Barr virus
  nuclear antigen 3A, 3B, and 3C homologues encoded by the rhesus
  lymphocryptovirus.";

```

J. Virol. 74:5921-5932(2000).
 [4] SEQUENCE FROM N.A.
 RC STRAIN=LCL8664;
 RX MEDLINE=20440633; PubMed=10970361;
 RA Rao P., Jiang H., Wang F.;
 RT "Cloning of the rhesus lymphocryptovirus viral capsid antigen and
 PT Epstein-Barr virus-encoded small RNA homologues and use in diagnosis
 RT of acute and persistent infections.";
 RL J. Clin. Microbiol. 38:3219-3225(2000).
 [5] SEQUENCE FROM N.A.
 RC STRAIN=LCL8664;
 RX MEDLINE=21602573; PubMed=11739708;
 RA Rivallier P., Jiang H., Cho Y.-G., Quink C., Wang F.;
 RT "Complete Nucleotide Sequence of the Rhesus Lymphocryptovirus: Genetic
 RT Validation for an Epstein-Barr Virus Animal Model.";
 RL J. Virol. 76:421-426(2002).
 [6] SEQUENCE FROM N.A.
 RC STRAIN=LCL8664;
 RX Moghaddam A., Koch J., Amis B., Wang F.;
 RA SubMITTED (MAR-1997) to the EMBL/GenBank/DBJ databases.
 [7] SEQUENCE FROM N.A.
 RC STRAIN=LCL8664;
 RX Moghaddam A., Amis B., Wang F.;
 RA SubMITTED (MAR-1997) to the EMBL/GenBank/DBJ databases.
 [8] SEQUENCE FROM N.A.
 RC STRAIN=LCL8664;
 RX Rivallier P., Quink C., Wang F.;
 RA SubMITTED (MAY-1999) to the EMBL/GenBank/DBJ databases.
 [9] SEQUENCE FROM N.A.
 RC STRAIN=LCL8664;
 RX Jiang H., Wang F.;
 RA SubMITTED (JUN-1999) to the EMBL/GenBank/DBJ databases.
 [10] SEQUENCE FROM N.A.
 RC STRAIN=LCL8664;
 RX Rao P.V., Jiang H., Wang F.;
 RA SubMITTED (JAN-2000) to the EMBL/GenBank/DBJ databases.
 [11] SEQUENCE FROM N.A.
 RC STRAIN=LCL8664;
 RX Rivallier P., Jiang H., Cho Y.-G., Quink C., Wang F.;
 RA SubMITTED (JUN-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AY037858; AAK95476.1; -
 DR InterPro; IPR00234; Glycoprot_B.
 DR Pfam; PF00606; Glycoprotein_B; 1.
 DR ProDom; PD00693; Glycoprot_B; 1.
 SQ SEQUENCE 864 AA; 96646 MW; 89637230C484F5BD CRC64;

Query Match 39.98; Score 1812; DB 12; Length 864;
 Best Local Similarity 45.44; Pred. No. 1.4e-105;
 Matches 366; Conservative 148; Mismatches 242; Indels 50; Gaps 14;

QY 81 STFFRYVCSAGVGDFVRFOTDHCPCDA-SDMVHSEGILLIYKQNIIPFMRVRYKRVV 139
 DB 44 TNFFRVCELSSHGDLFRFSDIQCPSSGTQEKHTEGLLVAFKDNIIIPYSFKVRSYTKIV 103
 QY 140 TTSTVYNGIYSDSTNTOHTFYKSIETPWEKMTDIYQCFNSRLNTGNNLLTVYVDRDIN 199
 DB 104 TNLTVNGVYADAVTNRHEMFVSVEHETQMDSIYQCHNAVMTKDGVERVYVNDGNTN 163
 QY 200 MTVPQLQVDGVTDPVKEVSGOPELYLEPGHFWGYSRRRTTVNCELMDMFARSNPPDFV 259
 DB 164 ITVNLRTSGLVNGVRYASQTQVHDAPGILMTYRTTRTTVNCVVTDMMAKSNPFFV 223
 QY 260 TATGDTVEMSPFWSGEGDDHKNKHEKPFWSVNNYKVVDYQNRGTVPGLGKTRIFLDREE 319
 DB 224 TTVGQTAEMSPFFYDGS--TNKETFNERAGSFHVENYKIVDYDNRGTVPNGERRAFLDKGT 281

QY 320 YTLSEKHLKN-MSYCPLTLWKAFYNGIOTHSRSGSYHFVANDITASFTTSKEDKSENTT 378
 DB 282 YTLSSWK--LENRTAYCPLQHWKTFESTIATETVRSFTQEGTSFVNTMTVGMALPGA 339
 QY 379 YHCLNEEIKAEIEKAYKNVSTHSK-YGDLKVPKTDGGLVYVWQPLIQNRLLDANKLNN 437
 DB 340 FKIEDQVNTKHEKYEALQDRYVKQENITFTLTSGLLLAWPLTPRSLASVKLTEL 399
 QY 438 ETYS-----RRSROAE--STT-----DPMEMTNGAGGEY 467
 DB 400 TTAAPPPSPPTPTSTASRSTAAAI VRRRRRQAEGNNTTPAPPKPATPAPGNSLG--- 456
 QY 468 SSENSITVAQVAYDNLRINNIILEDLSKAWCRHQHRAALVNNELSKINPTSVSMIY 527
 DB 457 -TLNPPATIQPAYDSLRQRINRMGLDLARAWCLQKQNMVLRRLTKINPTVMSGIY 515
 QY 528 NRPVSARKIGDVISVSNCIWVQTSVLSLHKSRLLSASDEKCFSPRPVTFKFWNDSTIYK 587
 DB 516 GRPVAARKLDVISVSCQVPVQATVTLKSMR-VGSESMCYSRPLVSFSFINDTKTYE 574
 QY 588 GOLGVNNEILLTTLTYLETQENTTEYFOAKTDMYIYKNYEHKTVPLSSITLDTFIALN 647
 DB 575 GOLGTDNEIFLTKMTEMCLASSQYVFQSGNEIHVVHYDQHPKTIELDGIATLQTFIALN 634
 QY 648 FTULENVDFKVELYTRDEKRLSNVEDIETMREYNYAQRVSGRLKDLID-LSHNRNQF 706
 DB 635 TSLIENIDFASLELYSRDEQRASNVFDEGI FREYNFOAQNTAGRLKDLNAYSNRNQF 694
 QY 707 VDAFGSLMDDLGAVGOTVNVNAVSGVATLFSIVTGFINKPFGGMLMIIVIGLVFAI 766
 DB 695 VDGLGELMDSLGSGVQVITNLVSTVGGLFSLNLSGFSIFPKPFGGMLTLVLVAGVVILV 754
 QY 767 YFLTKTKTYETAPIKMIYPEIDKLKER----EGKSE-IAPISSEELERIVLAMHIHQON 821
 DB 755 IITLTRTROMSQPVQMLYPGIDELARQHAGMGQGGQGNITPISKPELDAIMLA--LHEQN 812
 QY 822 SHMETKTRKPKDSILTRAQNMRLKR 847
 DB 813 QEQRAAQRAATKPSVASRALQAARD 838

RESULT 21
 Q98336 PRELIMINARY; PRT; 849 AA.
 ID Q98336; AC Q98336; DT 01-FEB-1997 (TrEMBLrel. 02, Created)
 DT 01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
 DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
 DE Glycoprotein B precursor homolog.
 GN GB OR GAMMAHV. ORF8 OR 8.
 OS Murid herpesvirus 4.
 OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
 OC Gammaherpesvirinae.
 OX NCBI_TaxID=33708;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=94365953; PubMed=8083987;
 RA Stewart J.P., Jianjua N.J., Sunil-Chandra N.P., Nash A.A., Arrand J.R.;
 RT "Characterization of murine gammaherpesvirus 68 glycoprotein B (gB)
 RT homolog: similarity to Epstein-Barr virus gB (gp110).";
 RL J. Virol. 68:6496-6504(1994).
 [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=WUMS, and G2.4;
 RX MEDLINE=97366649; PubMed=9223479;
 RA Virgin H.W. IV, Latraille P., Wamsley P., Hallsworth K., Weck K.E.,
 RA Dal Canto A.J., Speck S.H.;
 RT "Complete sequence and genomic analysis of murine gammaherpesvirus
 RT 68.";
 RL J. Virol. 71:5894-5904(1997).
 [3]
 RP SEQUENCE FROM N.A.

RT "Cloning and characterization of rhesus cytomegalovirus glycoprotein
B";
RL J. Gen. Virol. 78:2009-2013(1997).
DR EMBL: U59238; AAB70023.1; - - - - -
DR InterPro: IPR000234; Glycoprot_B.
DR InterPro: IPR000847; HTH_LysR.
DR Pfam: PF00606; Glycoprotein_B; 1.
DR ProDom: PD000693; Glycoprot_B; 1.
DR PROSITE: PS00044; HTH_LYSR_FAMILY; 1.
SQ SEQUENCE 854 AA; 97859 MW; 7AAEACAEADD9F69D1 CRC64;

Query Match 24.8%; Score 1126; DB 12; Length 854;
Best Local Similarity 31.5%; Pred. No. 2e-62;
Matches 281; Conservative 172; Mismatches 336; Indels 102; Gaps 27;

QY 20 VALYSLSLTAETGVS---PPNTATWSTESPLTGHGTHDSHGCRGNENRDSERQNKVI 76
DB 15 VVYVSIASSSTGTASGAVTPTSPIENTTGPLI-----ENTLTRTHEYFKINM 61
QY 77 YGSPSTFPYRVCSAGVGDFVRFQTDHVC---PDASDMVHSEGIILLIYKONIIPFMRV 132
DB 62 ---SKFPYRVCSMAQGTDLTLPFQINCDSEKPKEDF---EEGMVYKRDIDRPTFKV 115
QY 133 RYKRVVTTSTVYNGIYSDSINTQHTFYKSIPEWETKMDTIYQCFNSLRNLNTGGLTY 192
DB 116 HMYQKILTPQSYSTIRENHLGLFSGOHLAVPMWEVHYINKLRNCYNSVVRNVAG--ATY 173
QY 193 VD--RDD--INMTVFLOPQDGVTPDKRVYGSQBELVLEPGWFGVSRRTTNCCLMDNPA 249
DB 174 VNYHRDSYVNETMLLVEDDFSNTHSRFVTVKELHFKGSTW--LYTTSNCVNCMVTVTTA 232
QY 250 RSNPPDFEFTVATGTQVEMSPWSEDD--HENKMEKPFVSVINNYKVYQNRGTVP- 307
DB 233 RSKYDFDFVSDGKVDIDISPYNSGNKHFGENRDK---FSVRKNYSIAIYGRDNAP 289
QY 308 ---LGKTRIFLDREYTLSEKHLKNMSYCPLTLKAFYNGIOTERHSGSYHFVANDIT 362
DB 290 VAHPLVG---PFRPDVLMWDIVVEAANTCEYTFWEQSEBTIRSEADDIYHYTSSMT 345
QY 363 ASFTTSKEDMKFNITYHCL-----NEETKAEIKKYAKVNSHSGYGLKPKTDGGYLL 418
DB 346 ATFLSKBELNESDFSCIKDKANEOLQIFNTSY---NETYQSGVNSVYETTGGLIV 402
QY 419 VMOPLIONRLDKN---KLNNETYSRRSRQAESTTDPMMEMTNGAGGESVP--NSI 473
DB 403 FPLPKYKSLMEELAVAYNTNSSTRKSTDSASD-----SNKISEEVLKSI 452
QY 474 TVAQVQYAYDNLRIINNILEDLSKAWCREQHRALVWNELSKINPTSVMSIYNRPVSA 533
DB 453 VYALQYIYDILRNINRALRQIAEAWCKDQKRTLEVEFKELSKINPSAMLSAIYDKPIAA 512
QY 534 KRIGDIVSVNCIIVDQTSVLHSLRLLSASDEKCFRPPVTFKPMNDSTIYKQGLVN 593
DB 513 RFGVDVSLAKCVEVDONSVKVLRDMR--TKESGVCYSRPVVLYTFKNSHVQYQGLGEY 570
QY 594 NEILLTITTYLETCOENTYEFQAKTDMYIKNYEHLKTVPLSSITLDTFTALNFTLEN 653
DB 571 NEILLGRHTEACEYPSLKIYAGNSSYEVVDLYKRMIPLOSISTVDMISLSDIDPLEN 630
QY 654 VDFKVIELYTRDEKLSNVFDIETMFREYNYAQRVSGRLKDLDLSTNRNQVDAFGSL 713
DB 631 TDFKALELYSEDELSSNVFDELDIMREFNTYKQRMHVEGKVPKVPK---YLRGLDM 687
QY 714 MDGLGAGGTVNVANSGVATLFSSTVTFINIKNPPFG--MLMIIVIGVLFAYFLPT 770
DB 688 MSGLSAGKALGAVIAGVGAVASVEGVGFIKNPFGSFVILPFLAVGLVILYIMRQ 747
QY 771 KKT--KIVE-----TAPIKMIYPEIDKLEREGKSEIAP-----ISEELERIVLAWH 816
DB 748 KRAYEKPEHFPYVVPPTTVKEAPPSEYQOYENIKEAAPATKFEFLEEAYQMLLA-- 805
QY 817 IHQNSHMETKTRDKPKD-----SILTRAQNLRLKRSYSLNKLNAESV 859

DB 806 LQKLDQEKRRKAEADDEDFAHQAGFLDLRLN--RRRGYQKIQNEYEV 854

RESULT 28
Q69169 PRELIMINARY; PRT; 906 AA.
ID Q69169
AC Q69169;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Glycoprotein B.
OS Human cytomegalovirus.
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Betaherpesvirinae; Cytomegalovirus.
OX NCBI_TaxID=10359;
RN [1]
RC SEQUENCE FROM N.A.
RP STRAIN=C325A;
RX MEDLINE=92230241; PubMed=1314465;
RA Chou S.;
RT "Comparative analysis of sequence variation in gp116 and gp55 components of glycoprotein B of human cytomegalovirus.";
RL Virology 188:388-390(1992).
DR EMBL: M60927; AAA45926.1; - - -
DR InterPro: IPR000234; Glycoprot_B.
DR InterPro: IPR000847; HTH_LysR.
DR Pfam: PF00606; Glycoprotein_B; 1.
DR ProDom: PD000693; Glycoprot_B; 1.
DR PROSITE: PS00044; HTH_LYSR_FAMILY; 1.
SQ SEQUENCE 906 AA; 101833 MW; 523C585377AAC9F2 CRC64;

Query Match 24.7%; Score 1123.5; DB 12; Length 906;
Best Local Similarity 30.9%; Pred. No. 3.1e-62;
Matches 286; Conservative 182; Mismatches 353; Indels 105; Gaps 24;

QY 14 LWLYQVALYSLSLTAETG--VTSPPNTATWSTESPLTGHYTHDSS-----HGERG 62
DB 5 IWCL--VVCNLCIVCLGAAVSSSTSHATST-----HNGSHTSRTTSQATRSVSSQHV 56
QY 63 NENRDEEONKNTYG-----SPSTFPYRVCSAGVGDFVRFQTDHVC-----PD 107
DB 57 TSSEAVSHRANETIYNTLTLYGVDVGVNTTKYPRVCSMAQGTDLIRFERNIVCTSMKPI 116
QY 108 ASDMVHSEGIILLIYKONIIPFMRVRYKRVKVVTTSTVYNGIYSDSINTQHTFYKSIPEWE 167
DB 117 NEDL--DEGIMVYKRNIVAHFTKRVYQVLTFRYSYAIHTTYLLGNSNTEYVAPPWE 174
QY 168 TEKMDTIYQCFNSLRNLNTYVDODDI--NMTVFLOPQDGVTPDKRVYGSQPELYLE 226
DB 175 IHHNSHSQCYSSYSRVIAGTVFVAYHRDSYENKTMQLMPDDYNTSHSTRVTVKDWHS 234
QY 227 PGWFWGYSRRRTTNCCLMDMFARSNPPDFVTFATGDTVEMSPFWSEGDHKNHMKP 286
DB 235 RGSTW--LYRETCLNLCMVTTITARSKYPHFFATSTGDVVDISPFGYNTNRNASYFGENA 293
QY 287 WFSVINNYKVVDYQNRGTVPGLKTR--IFLDREYTLSE--KHLKNMSVCLTLWAKFY 343
DB 294 DKFFIFPNYITVDFGRPSALETHRSVAFLERADSVISWDIODEKNT--CQLTFWEASE 352
QY 344 NGIOTERHSGSYHFVANDITASFTTSKEDMKFNITYHCLNEETKAEIKKY--AKVNSTHS 402
DB 353 RTIRSEADESYHFSACKMTATFLSKQEVNMSDALDCVDEAKINKLOQIFNTSYNTYE 412
QY 403 KYGDKLYFKTDGGYLLVWQPLIQNRLDANKLNNETYSRRSRQAESTTDPMMEMTNG 462
DB 413 KYGNVSVEFTTGGVFWQGIKQKSLVELERLANRSSNLNLTNRKST-----NGNN 465
QY 463 AG--GEVSSSENSITVAQVAYDNLRIINNILEDLSKAWCREQHRALVWNELSKINPT 520
DB 466 ATHLSNMESVHNLVYALQFTYDTLGRYINRALAQIAEAWCVQDRRSLEYFRELKINPS 525
QY 521 SVMSIYNRPVSARKIGDIVSVNCIIVDQTSVLHSLRLLSASDEKCFRPPVTFKFM 580


```
DR InterPro; IPR000847; HTH_LysR.
DR Pfam; PF00606; Glycoprotein B; 1.
DR ProDom; PD000693; Glycoprot B; 1.
DR PROSITE; PS00044; HTH_LysR_FAMILY; 1.
SQ SEQUENCE 905 AA; 101916 MW; 90BEE74D7581EFF1 CRC64;

Query Match      24.6%; Score 1119; DB 12; Length 905;
Best Local Similarity 30.9%; Pred. No. 5.9e-62;
Matches 286; Conservative 186; Mismatches 349; Indels 104; Gaps 25;

QY 14 LWLYQVALYSIAETG--VTSPTATWSTESPLTGHYGDHDS-----HGERGNE 65
DB 5 IWCL--VVCNLCI VCLGAUVSSSTSRAT-----STHNGHTSHSTSAQTSRISQVSTSS 58
QY 66 NRDSSEQNKNIYG-----SPSTPPYRVCSGASGVDFRFTDHCV-----PDASD 110
DB 59 EAVSHRANETIYNTILKYGDVVGVNTTKPYRVCSMAQGTDLIRDRNIVCTPMKPINED 118
QY 111 MVHSGILLIYKONIPMFRVRKYKVVTTSTVYNGIYSDSITNQHTFYKSIETPWEYTEK 170
DB 119 L--DEGIMVYKRNIVAHTEKRVVYQKVLTPRRSVAYIHTTYLLGSGNTEYVAPPWWEIHH 176
QY 171 MDTIYQCNLSRLNTGNNLLTVDRDDI--NMVFLQPDVGVTPDKVRYGSGQPELYLEPGW 229
DB 177 INSHSQCYSSYSRVIAGTIVFAVHRDSYENKTMQLMDDYDYNTHSTRYVTVKDDQWHSRGS 236
QY 230 FWGYSRRRTTVMCELMDFARNSPPDFVTATGTVEMSPFWSGEGDDHKNMHEKPFV 289
DB 237 TW-LYRETCNLNCMTIITARSKYPIYHPFATSTGDDVDLSPFYNTNRRNASVFGENADKF 295
QY 290 SVINNYKVDVQNRGTPLGKTRI--FLDREYTLSE--KHLKNMSYCPLTLWKAFYNGI 346
DB 296 FIFPNYTVSDPGRNSALETHRLVAFLERADSVISWDIQDEKNTV--COLTFWEASERTI 354
QY 347 QTEHSGSYHFVANDITASFTTSKEDMKENITYHCLNEEIKABIEKKY--AKVNSTHSYK 405
DB 355 RSEAESYHFSKATATATFLSKQKQVNMSSDALDCVRDEALNKLQOI FNASNYQTYEYK 414
QY 406 DLKYEKTDGGLYLVNQPIQNRLLDAKNKLN---NETYSRRSRQAESTTDPMMEMTGN 461
DB 415 NVSVETGGVVVFGQIKQKSLVLERLANRSSLNLTHS--RTRSTGTGNTVTHL----- 468
QY 462 GAGGYSSENSITVAQVAYDNLRIINILEDLSKAWCREQHRRAALVWNELSKINPTS 521
DB 469 ---SNMDSVHNLVYAQLQFTYDLRGYNRALQTIAEAWCVDQRTLEVFKELSKINPSA 525
QY 522 VMSMYNRPVSKRIGDVISVNCIVDQTSVLSHKSRLLSASDEKCFSPRPVTFKPMN 581
DB 526 ILSAIYNKPIAARFNGDVLGASCVTINQTSYKVLDRMN--VKESPGRCYSRPVVFNFAN 584
QY 582 DSTIYKGLGVNNEILLTTTLETQENTYFYQAQKTDWYIKNYEHLKTVPLSSITLD 641
DB 585 SSVYQYGLGEDNELLGNHRTTEECQLPSLKIF IAGNSAYEYVDYFLFRMIDLSSISYVD 644
QY 642 TFIALNFTLLENVDFKVIETRYDEKRLSNVFDIETMFREYNYAQRVSGRLKOLLDLST 701
DB 645 SMIALDIDPLENTDFRVLLEYLSQKELRSSNVFDLEEIMREFNYSYKQRYKVEDKVD--- 701
QY 702 NRNQVDAFGSLMDDLGAQGQTVVNAVSGVATLFFSSIVTGFINFINKNPFGLMLIIVVI 760
DB 702 PLPPYLKGLDMLSGLGAGKAVGVAIGAVGAVASVVEGVATFLKNPFGAFTIILVAIA 761
QY 762 VLFAIYFLTKTKIYETAPIKMIYPEI-----DKLKE-----R 794
DB 762 VVILIYLTQRRLCMQPLQNLFYLVLSADGTTVTSGNTKDTLSQAQPSVEESVYNSGR 821
QY 795 EG-----KSEIAPISEELERIVLAM-----HIHQONS-----HMETKTR--KDPKDS 835
DB 822 KGGPPSPSDASTAAPPYNEQAYQMLLALVRLDARQAQQGTDSLDQGTGTQDKGQKPN 881
QY 836 ILTRAQNMLRKRGVSNLKNAESVE 860
DB 882 LLDLRLH---RKNGYRHLKDSDEE 903
```

```
RC STRAIN=C076A;
RX MEDLINE=92230241; PubMed=13114465;
RA Chou S.;
RT "Comparative analysis of sequence variation in gp16 and gp55
RL components of glycoprotein B of human cytomegalovirus.";
RL Virology 188:388-390(1992).
DR EMBL; M85228; AAA45934.1; -.
DR InterPro; IPR000234; Glycoprot_B.
DR InterPro; IPR000847; HTH_LysR.
DR Pfam; PF00606; Glycoprotein_B; 1.
DR ProDom; PD00693; Glycoprot_B; 1.
DR ProSITE; PS00044; HTH_LYSR_FAMILY; 1.
SQ SEQUENCE 905 AA; 101868 MW; AFDA168807436890 CRC64;

Query Match 24.4%; Score 1110; DB 12; Length 905;
Best Local Similarity 30.7%; Pred. No. 2.2e-61;
Matches 280; Conservative 188; Mismatches 358; Indels 85; Gaps 23;

14 LWLYQVALYSLSLAETGVTPPN-----TATWSTESPLTGHVGDTHDSHGGRGNENR 67
14 LCVCLGAVVSSSTSHATSSAHNGSHTSRTTSAQTRSVSSQHV-TSSEAVSHRANETIY 72
68 DSEQNKNHYGSPST-FPYRVCSAGVGDVFRFQTDHVC-----PDASDMVHSEGIILYK 122
73 NNTLKYGDVVGVTTKYPYRVCSMAQGTDLIRFERNIVCTPMKPINEDL--DSGIMVYK 130
123 ONIIPMFRVRKYRKVVTSTVNGIYSDSIHQHTFYKIEPWETEKMDTIYQCNSLR 182
131 RNIVAHTFKVRVYQKLTFRSAYIHTTYLLGNTSEYVAPPWWEIHHNRHSQCYSYS 190
183 LNTGGNLLTVVDRDDI-NMTVFLQPDVGVTPDKVRKGSQPELYLEPGWFGSVRRRTTN 241
191 RVTAGTVFVAHYRDSYENKTMQLMDDYSNTHSTRVTVKQDQWHSRGSTW-LYRETCNLN 249
242 CELMDMPARSNPDPFVATGTVEMSPFSGEDDHENKMKHKPMFVSVINNYKVY-DY 300
250 CMVTITTKRSKYFHFATSTGVDVDSIPFYNGTNRNTSYFGENADKFFIFPNYTIVSDF 309
301 QNRGTVP-LKTRIFLDREYITLSE--KHLKNMSYCBLLTWKAFYNGIQTEHSGSVHFA 358
310 GRANSAPETHRLVAFLERADSVISWDIQDEKNVT-CQLTFWEASERTIRSEADSYHFS 368
359 NDITASPTTKEDMKENTTYHCLNEBEIKAEIKKY-AKYNSTHSKYGDLYKFTDGGLY 417
369 AKMTATFLSKQEVNMSDPLVDVCDQALNKLQIFNASNQYIEKYGNVSVETTGGLV 428
418 LVNQPLIQNLLDAKKNLANE--TYSRRSRQAEISTTDPMMWMTGNGAGGEYSSENSITV 475
429 VFQGIQKSLLELERLANGSGVNSTRTTKRSTGNTTTLSES-----ESVRNVLY 479
476 AQVQAYDNRIIRINNILEDLKAWCREQHRALVNNELSKINPTSVMSMIYRNPVSAKR 535
480 AOLQFTYDTRLRSINRALAQIAEAWCYDQRTLEVFKELSKINPSAILSAYINKPTAARF 539
536 IGVDSIVSNVVDQTSVLSHLKRLLSASDEKCFSPRPVTFKFMNDSTTYKLGQVWNE 595
540 MGDVLGLASCUTINQTSVKVLRDMN-VKESPGRCYSRPVVIFNVFNSVYVOYQLGEDNE 598
596 ILTTTYLETQCENTEYFQAKTDMYIKNYEHLKTVPLSSITLDTFFIALNFTLLENVD 655
599 ILLGNHRTBECQFPLSKIFIAAGNSAYEYDYLFRKMDLSSISTVDSMIALDIDPLENTD 658
656 FKVIELYTRBKLSNVFDIETMFREYNYAQRVSGRLKOLLDDLSTNRNOFVDAFGLMD 715
659 FRVLELYSQKELSSNVFDEEIMREFNYSYKQRYKVEDKVVD---PLPPYLKGLDGLMS 715
716 DLGAVGTVNVNAGSVATLFSSTVGTGFINKNPFPGGLMLIIIVIGVLFAYFLTKTKI 775
716 GLGAGKAVGVAICAVGASVVEGVATFLKNPFGAFTILVAIAVVIILYLYTRQR 775
776 YETAPIKMIYPEI-----DKLKE-----REG-----KSEI 800
776 LCMQPLQNLFPYLVSADGTTVTSNTKDTSLQAPPSVEESVYNSGRKGPSPSSDASTAA 835
801 APISBEELERIVLAM-----HIHQONS-----HMETKTR-KDPKDSILTRAQNMRLKRS 849
836 PPTNEAYQMLLALARLDAEORAQNGQTDSDGQTGDTQKQKPNLDRLRH---RKNG 892
850 YSNLKNASVE 860
893 YRHLKSDSEE 903

RESULT 33
Q992K7 PRELIMINARY; PRT; 340 AA.
AC Q992K7;
DT 01-JUN-2001 (TReMBLrel. 17, Created)
DT 01-JUN-2001 (TReMBLrel. 17, Last sequence update)
DT 01-OCT-2001 (TReMBLrel. 18, Last annotation update)
DE Glycoprotein B (Fragment).
GN GB.
OS Bovine lymphotropic herpesvirus.
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC unclassified Herpesviridae.
OX NCBI_Taxid=69955;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21212029; PubMed=11311431;
RA Chmielewicz B., Goltz M., Ehlers B.;
RT "Detection and multigenic characterization of a novel gammaherpesvirus
RT in goats.";
RL Virus Res. 75:87-94(2001).
DR EMBL; AF327832; AAK28845.1; -.
DR InterPro; IPR000234; Glycoprot_B.
DR Pfam; PF00606; Glycoprotein_B; 1.
DR ProDom; PD00693; Glycoprot_B; 1.
FT NON TER 1
FT NON TER 340
SQ SEQUENCE 340 AA; 38661 MW; C1287A7802AD81F5 CRC64;

Query Match 24.1%; Score 1097; DB 12; Length 340;
Best Local Similarity 59.0%; Pred. No. 3.7e-61;
Matches 203; Conservative 76; Mismatches 55; Indels 10; Gaps 4;

526 IYNRPVSAKRIGDVSNSNCIVVDQTSVLSHLKRLLSASDEKCFSPRPVTFKFMNDSTI 585
3 IYKGPVSARKLGDALSVCQVVDQSSVLSLHKSMT--QSAGDCYSRPLVTFKFLADTNV 60
586 YKQGLGVNNEILLTTLTYLETQCENTEYFQAKTDMYIKNYEHLKTVPLSSITLDTFFIA 645
61 YRGLGVSNELILLTTSVAEACHDGHYFQGAHMYKKNYEHTETIPLSAVSTLNTFIV 120
646 LNFTLLENVDPKVIELYTRDEKLSNVFDIETMFREYNYAQRVSGRLKOLLDDLSTNRN 705
121 LNLTLLENDFPHVIELYSREERKLSVFDIETMFREYNYTHRMSSGIKKDLNLDLATNRN 180
706 FVDAFGLMDDLGAVGTVNVNAGSVATLFSSTVGTGFINKNPFPGGLMLIIIVIGVLEA 765
181 FIDVFGSLMDDLGIGIKTVINAGSVATWTFESIVTGIVNFIKNPFGGLVFGIILVII 240
766 IYFLTKTKIYETAPIKMIYPEIDKLKREGK---SEIAPISBEELERIVLAMH-IHQON 821
241 VFMLNKRKNFEQDPVKVIYVDIQIKREQQEKDPENKQVPISKEELDKIMLAHMYVQKN 300
822 SHMETKTRKPKDSILTRAQNMRLKRSQVSNLKNASVEMLNTL 865
301 SEHE-----KNENNTLMDKAKNVVRKSGYRPLNKSESIEMLNIV 340

RESULT 34
Q39989 PRELIMINARY; PRT; 822 AA.
AC Q39989;
DT 01-JAN-1998 (TReMBLrel. 05, Created)
DT 01-JAN-1998 (TReMBLrel. 05, Last sequence update)
```

01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Glycoprotein B.
OS Human herpesvirus 7.
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Betaherpesvirinae.
OX NCBI_TaxID=10372;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=AL;
RX MEDLINE=97413886; PubMed=9268767;
RA Secchiero P., Berneman Z.N., Sun D., Nicholas J., Reitz M.S. Jr.;
RT "Identification of envelope glycoproteins H and B homologues of human
herpesvirus 7.";
RL Intervirology 40:22-32(1997).
DR EMBL; AF007829; AAB63200.1; -
DR InterPro; IPR00234; Glycoprot B.
DR Pfam; PF00606; Glycoprotein B_1.
DR ProDom; PD000693; Glycoprot_B; 1.
SQ SEQUENCE 822 AA; 93176 MW; 052F7E23AD257D1 CRC64;

Query Match 24.0%; Score 1090.5; DB 12; Length 822;
Best Local Similarity 33.7%; Pred. No. 3.2e-60;
Matches 273; Conservative 146; Mismatches 314; Indels 77; Gaps 24;

Qy 78 GSSTPTPYRVCASGVGDVRFOTDHVCPD-ASDMVHSGILLIYKONIIPMFVRKYR 136
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
28 GHNQHLPRFCISATGDLVRFDREVSCASYSNIKTEGILIIYKTIEAHTFSVRTFK 87
Qy 137 KVTTSIVNGI----YSDSITNQHFTYKSIEPWETEKMDTIYQCENSLRLNTG--GNLL 190
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
88 KELTFQTTRDVGTGYFLDRTVT--TLPMPIE--EVHMVNTEARCLSSIVSKRSSEBYV 143
Qy 191 TYVDRRDINNVTFLQPVGDTVPDKRYGSOPELYLEPGFMWGSYRRRTTVNCCLMDPEAR 250
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
144 AYHKDEVNKTLDLIPLNFKSDTVRIYITKPFFLANGPLM-FYSTSGINCIVTDCIAK 202
Qy 251 SNPPFPFVFATGDTVEMGFWSGEDDHE-NKMKEKPFVSVINNY---KVVDYQNRGTV 306
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
203 TKYPDFDFALLSETVEGSPFYINGINSKTFNEPTEKILF--RNNYTMKTFDDGSKGNF 259
Qy 307 PLCKTRIFLDREBYTISWEKHLKMSYCPLTLWKAFYNGIQTEHSGSHVFVANDITASFT 366
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
260 VTUKMAFLKGNITPSWEQNBESICLLKHMTTIPHALRAENANSFHFIQAELTASFV 319
Qy 367 TSXEDMEFNNTYHYCLNEBEIKAEIKKY-AKNVNSTHSYGDLKYFKTDGGLVWQPLIQ 425
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
320 TGKSNYTLSDSKYNCSINYTSILDYIOTYNSHDKNGSYEIFKTEGDLILIWQPLIQ 379
Qy 426 NRLLDAKKLNNTYSRRRRQRQAESTTDPMMENTGNAGGEYSSENSITVAQVOYAYDNL 485
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
380 RKLTVLEN-FSNSAKRRKR-----ELETKNDIVYVQLQLYDYL 418
Qy 486 RIPIINILEDLSKAWCEQHRALVNWELSKNTPTSVMNMVNRPVSAKRIGDIVISUNC 545
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
419 KOYINTALGKAIAEWCLDQKRTITVLHELKSIKSPGIIISAVYKGPKMSAKLIGDLVASKC 478
Qy 546 IVVDQTSVSLHSLRLLSASDE---KCFSRPPVTKFMNDS-TIYKGGOLGVNNEILLTTT 601
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
479 IEVNOTSVQLHKSMLTKUSSVDALRCYSRPLTYTSFANSSKETYGQGLDNEILLGNH 538
Qy 602 YLET-C-QENTYYFOAKTDMYIKNYEHUKTVPLSSITLTDFIALNFLTENVDKFVIE 660
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
539 RTECEQSNTKI FUSGKF-AHFIDKYTYVNSSLITEIBALDAFVLDINDPLENVDFILLE 597
Qy 661 LYTRDEKRLSNVEDIETMFEREYNYAQRVSGLRKOLLDSNRNOFVDAFGSLMDDLGA 720
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
598 LYTKELSKANVPDELTELIREVNSYK--SALHHIETKIATVPTYIGSIDTFFKGLGAL 654
Qy 721 GQTVNAVSGV-ATLFSSIVTGFINFKPFGGMIMIIVIGLVFAIYFLTKTIYETA 779
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
655 G-LGLGAVLGVTAGALGDVVNGVFSFLKMPFGALTILTULGVGLVIFLFLRHKRLAQT 713
Qy 780 PIKMIYPEIDKLKER--EGKSGIAPISEBELERIVLAMHIHQONGHMETKTRKD--PKDS 835

```

QY 410 FKTDGGLYWOPLIONRLDAKLNKNETYSRRSRAQESTTDPMMEMTNGAGGEYSS 469
Db 407 YETTGILYWOPLIVKRSIKELKEFINQHTQRYKR-----BIDG 446
QY 470 ENSITVAQVOYAYDNLRIRINNILEDLSKAWCREQRAALVMNLSKINPTSVMSMIYNR 529
Db 447 SDSLVYASLOMYDALREYINAGFAQIAEAWCEDQKRTNEVLSELAKISPSNVMSVIYDK 506
QY 530 PVSAKRIGDVISVNCIYVDOOTSLSLHKSRLLI---SASDEKCFSPRPVTFKFWNDSTI- 585
Db 507 SLSAKLVGDAISVSSCVNWNQSTVKVHKDMRIYANGTANRETCSRPVVTFFESNNNSVQ 566
QY 586 -YKGLGQWNEILLTLYLTCENTEYFQAKTDMYIKYNEHLKTVPLSSITLDTLTFI 644
Db 567 QFTGQLGRNEILLGTTRVCEKERNVKKVFPAGKEAYFFNYIYTKTVNISDINVDTFI 626
QY 645 ALNFTLENDVDFKVIELYLDEKRLSNVFDIETMFREYNYAQRVSGLRKDLDDLSTNRN 704
Db 627 HLNKPLENTDFEVLRYMSKNELAQAANIFDLESILRDISYRSALYNIES---RIAPRKP 683
QY 705 QFVDAFGSLMDDIGAGVQTVNAVSGVAT-LFSSIVTGFNFINKPFGG---MLMIIVVI 760
Db 684 DYVSGVDSFLHALGIGAPGGLGALGMATGAVTDFLTGIFSPFKNPPGGLFMSLFFVLV 743
QY 761 GVLPAIYFLTKTKIYETAPIKMIY-----EIDKLERECK-----797
Db 744 FLIFSVIY--RQKNIY--TNVPGALFPYANSSGTVISNTHSYETNNKQEFENDRKPPTS 800
QY 798 ---SEIAPISEELERIVLAMHHQNSHMETKTRKDKDSILTRAQNM 843
Db 801 NAVSEGSANKYSQEDAVCMLEMAIKNLGDAYRRKNATKPSVLDKIRHL 849

RESULT 36
Q96910 PRELIMINARY; PRT; 822 AA.
AC Q96910:
DT 01-FEB-1997 (TrEMBLrel. 02, Created)
DT 01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Glycoprotein B.
GN U39.
OS Human herpesvirus 7.
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Betaherpesvirinae.
OX NCBI_TaxID=10372;
RN [1]
RP SEQUENCE FROM N.A.
RA Hata A., Mukai T., Isegawa Y., Yamanishi K.;
RT "Identification and analyses of glycoprotein B of human herpesvirus
7.";
RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL; D86577; BAAL13124.1; -
DR InterPro: IPR000234; Glycoprot B.
DR Pfam: PF00606; Glycoprotein_B_1.
DR ProDom: PD000693; Glycoprot_B; 1.
SQ SEQUENCE 822 AA; 93178 MW; E3BFAA858362A619 CRC64;

Query Match 23.9%; Score 1086.5; DB 12; Length 822;
Best Local Similarity 33.6%; Pred. No. 5.8e-60;
Matches 272; Conservative 146; Mismatches 315; Indels 77; Gaps 24;

QY 78 GSSESTFFRYVCSAGVGDFVRFQTDHVCDD-ASDMHSEGILLIYKQNIIPFVRVKYR 136
Db 28 GHNOHLPFRICSIATGDLVRDFREVSCASVGSNIKTTEGILIIYTKIEAHTFSVRTPK 87
QY 137 KVTTSIVYNGI----YSDSITQHTFYKSIETWEITEKMDTIYQCNSLRLNTG--GNLL 190
Db 88 KEUFTQTYRDVGTVFLDTRVT--TLPMPIE--EVHMVNTAEARCUSSISVKSSEEEYV 143
QY 191 TYVDRDDINMTVFIQPDVGVTPDVVKRYGSOPELYLEPGFWGSGYRRRTTNCVCLMDMFAR 250

```

```

Db 144 AVHKDEYVNTKLDLIPLNFKSDTVRRYITTKSPFLRNGPLW-FYSTSTSIINCIVTDCIAK 202
QY 251 SNPPPPFFVTATCDTVMSPFWSGEDDHE-NKMHKXPWFVSVINNY---KWDYQNRGTV 306
Db 203 TKYPFFDFALSTGETVEGSPFFYNGINSKTFNEPTEKILF---RNNTYMLTKTFDDGSKGNF 259
QY 307 PLGKTRIFLDREBYTLSEWEKHLKNMYSYCLTLTWKAFYNGIQTSHSGSYHYFVANDITASFT 366
Db 260 VILTKNAFLEKNGNTIFSNEVQNEESSICLLKHWMTIPHALRAENANSFPIAELTASFV 319
QY 367 TSKEDMKFEFTTYHCLNEEIKABIEKKY-AKVNSTHSKYGDLKYFKTDGGLYLVWQPLIQ 425
Db 320 TGSNVTLSDSKYNCINSVNTSILDEIYQTYQVNSHDKNGSYEIFKTEGDLILWQPLIQ 379
QY 426 NRLLDKAKLNNETYSRRSRAQESTTDPMMEMTNGAGGEYSSENSITVAQVOYAYDNL 485
Db 380 RKLTVLEN-FSNASRKRKR-----ELETNKDIIVVQLQYLDYDTL 418
QY 486 RIRINNILEDLSKAWCREQRAALVMNLSKINPTSVMSMIYNRVPSAKRIGDVISVNC 545
Db 419 KOYINTALGKLAEWCLDQKRTITVLHLSKISPSGIIISAVYKPKSAKILGDLVAVSKC 478
QY 546 IVDQTSVSLHKSRLRLSASDE---KCFSRPPVTFKFMNDS-TIYKGLQGVNNEILLTTT 601
Db 479 IEVNTQTSVOLHKSRLTKDSSYDTLRCYSRPLLTYSFANSKSETYLGQLGLDNEILLGNH 538
QY 602 YLETC-QENTEYFQAKTDMYIKYNEHLKTVPLSSITLDTFIALNFTLENVDFKVI 660
Db 539 RTEECQSNKIFLSGKF-AHIFKDYTVYNSSLITEALDAFVDINIDPLENADFTLLE 597
QY 661 LYTRDEKRLSNVFDIETMFREYNYAQRVSGLRKDLDDLSTNRNQFVDAFGSLMDDLGAV 720
Db 598 LYTQDELKANVDFLETLIREVNSYK---SALHHIETKIATVTPTVIGGIDTFFKGLGAL 654
QY 721 GQTVNAVSGV-ATLFSSIVTGFNFINKPFGGMLMIIVVIGLVFAIYFLTKTKIYETA 779
Db 655 G-LGLGAVLGTAGALGDVGVNGVFSPLKNPFGGALTILLTGLVIGLVFLFLRHKRLAQT 713
QY 780 PIKMIYPEIDKLER--EGKSEIAPISEELERIVLAMHHQNSHMETKTRD--PKDS 835
Db 714 PIDILFPYTSKTSNSTVLOATQSVQAQVKPLD-----SSPPYLKTNKDTPEQGD 762
QY 836 ILTRAQNMRLKRSYSNLKNABSVEMLNTL 865
Db 763 DIHTNE-----YSQV---EALXMKLXAI 782

RESULT 37
O56280 PRELIMINARY; PRT; 822 AA.
AC O56280:
DT 01-JUN-1998 (TrEMBLrel. 06, Created)
DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Envelope glycoprotein gB.
GN U39.
OS Human herpesvirus 7.
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Betaherpesvirinae.
OX NCBI_TaxID=10372;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=RR;
RX MEDLINE=98240941; PubMed=9581785;
RA Megaw A.G., Rapaport D., Avidor B., Frenkel N., Davison A.J.;
RT "The DNA sequence of the RK strain of human herpesvirus 7.";
RL Virology 244:119-132 (1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=RR;
RA Megaw A.G., Frenkel N., Davison A.J.;
RL Submitted (DEC-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF037218; AAC40753.1; -

```

DT	01-JUN-2001	(TREMELrel. 17, Created)
DT	01-JUN-2001	(TREMELrel. 17, Last sequence update)
DT	01-OCT-2001	(TREMELrel 18, Last annotation update)

DR ProDom; PD000693; Glycoprot B; 1.
 SQ SEQUENCE 859 AA; 97912 MW; 659C023F79CB5B2B CRC64;
 Query Match 23.3%; Score 1061; DB 12; Length 859;
 Best Local Similarity 30.9%; Pred. No. 2.5e-58;
 Matches 246; Conservative 177; Mismatches 304; Indels 74; Gaps 19;
 QY 11 VLALWLYQVALYSLSLA----ETGVTSPPN-----TATWSTESPLTGHVTHDSSHGE 60
 DB 12 VLLINYGVLSSNSTTATAPESVITGSSGTVISATPSTVSTAFVNITGNFQSEYTE 71
 QY 61 RGNENRDSBEQNKNIYGSSTPPYRVCSAGVGDFVRFQTDHVCPPD-ASDMVHSEGILL 119
 DB 72 ASDDEK-----YFPRVCNMAVGTDLYRFDNYITCNKYDTETQYSEGILL 115
 QY 120 IYKNIILPFMRVKYKVVTTSTVYNGIYSDSIHQHTFYKSIPEWETEKMDTIYOCFN 179
 DB 116 LFKKNIVPHFTFFVRYTYKELSFQTYRDVHVILVDRSSYKVPVDEAGYINLNGCFS 175
 QY 180 SLRL-NTGGNLLTYVDRDDINMTVFLQPDVGVTPDKRVGSQPELYLEPGWFGSVRRRT 238
 DB 176 AAEIRNOGINRVYHKDDNTNMRLYLXELFGSTINTRYISTPDTQFTYTHW-LYKSSS 234
 QY 239 TVNCELMDMFARSNPPDFVATGDTVEMSPWGSDEDDHENKMKHPWFSVNNYKV- 297
 DB 235 SINCIITDTMGKSDYPYDNFILTGESVEISPPFNCTS--KEVNEQMYFMSKRNVTML 292
 QY 298 --VDYQN--RGTVPLGKTRIFLDREBYTLSWEKHLKMSYCPLTLKAFYNGIQTEHSGS 353
 DB 293 EKLDPEKPKTIP--TIAFLQKGDTLFSWEVKQNTYHCKYTAWTKKHALRADMTNS 349
 QY 354 YHFVANDITASFTTSKEDMKFNFTYHCLNEEIKAEIEKKY-AKVNSTHSGYDGLKYFKT 412
 DB 350 YHFMKMDTATWTTKNTINLTGTEYECVKNDIEKYITDTFQSKYNNTHKNENYVET 409
 QY 413 DGGYLWQPLIQNRLLDANKLNNETYRSRRRQAEISTTDPMMWMTGNGAGGEYSSENS 472
 DB 410 TTGLILWQPIVRKSIKELKEFINETQOHTRYKR-----EIDGES 450
 QY 473 ITVAQVOYADNLRIRINNILEDLSKAWCREQRAALVWNLKINPTSVMNINRPVS 532
 DB 451 LVVASLQYMDALREYNAGFAQIAEAWCEDQKRTNEVLSELAKINPSNVMYIDKSL 510
 QY 533 AKRIGDIVISNCIVVDQTSVLSHKSRL--SASDEKCFSPPTVFKPMNDSTI--YK 587
 DB 511 AKLVGDALSVSSCVNVQSVKHKDMRIYANGTANRETCSRPVTFESSNNSVQOFT 570
 QY 588 GOLGVNNEILLTLYLETQENTYFYFOAKTDMYIYKNYEHKLTVPVLSITLDTFIALN 647
 DB 571 GOLGPRNEILLGTHRVKCEKERNVSVKVPFAGKEAYFFNYTYTKTVNISDINNVVDTFHLN 630
 QY 648 FTLEENVDFKVIELYTRDEKRLSNVFDIETMFREYNYAQRVSGRLKDLDLSTNRNPFV 707
 DB 631 IKPLENTDFEVLRMYSKNELAQANIFDLESLLIDINSYRSALYNIES---RIAPRKPDYV 687
 QY 708 DAFGLMDDLCAVGQTVVNAVSGVAT-LFSSIVTGFINKNPFEG--MLMIIVIGVL 763
 DB 688 SGVDSFHALGIGAPGGLGAALGATGAVTDLTGTFESFPKPNFGLFSLMFLVLLVFLI 747
 QY 764 FAIYFLTKTKIYETAPIKMIYP 786
 DB 748 FSVYY--RQKNII-TNPVGALFP 767
 RESULT 43
 Q8V613 PRELIMINARY; PRT; 860 AA.
 AC Q8V613
 DT 01-MAR-2002 (TrEMBLrel. 20, Created)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE Glycoprotein B.
 OS Porcine cytomegalovirus.

OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
 OC Betaherpesvirinae; Cytomegalovirus.
 OX NCBI_TaxID=109993;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=OF-1;
 RX MEDLINE=21011895; PubMed=11129643;
 RA Goltz M., Widen F., Banks M., Belak S., Ehlers B.;
 RA "Characterization of the DNA polymerase loci of porcine
 RT cytomegaloviruses from diverse geographic origins";
 RL Virus Genes 21:249-255 (2000).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=OF-1;
 RA Goltz M., Widen F., Banks M., Belak S., Ehlers B.;
 RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF268041; AAL47540.1; --
 DR InterPro; IPR000234; Glycoprot_B.
 DR Pfam; PF00606; Glycoprotein_B; 1.
 DR ProDom; PD000693; Glycoprot_B; 1.
 SQ SEQUENCE 860 AA; 97834 MW; E61F989F175BF433 CRC64;
 Query Match 23.3%; Score 1059; DB 12; Length 860;
 Best Local Similarity 29.1%; Pred. No. 3.3e-58;
 Matches 259; Conservative 194; Mismatches 330; Indels 106; Gaps 23;
 QY 11 VLALWLYQVALYSLSLA----ETGVTS--PPNTATWSTESP-----LTGHVGHDS 57
 DB 12 VLLINYGVLSSNSTTATAPESVITGSSGTVSTETATPEVSTAFVNITGNF----- 65
 QY 58 HGERGNENRDSBEQNKNIYGSSTPPYRVCSAGVGDFVRFQTDHVCPPD-ASDMVHSEG 116
 DB 66 ----SQEYTERASDDEK-----YFPRVCNMAVGTDLYRFDNYITCNKYDTETQYSEG 112
 QY 117 ILLIYKONIIPMRVRKYRKVVTTSTVYNGIYSDSIHQHTFYKSIPEWETEKMDTIYQ 176
 DB 113 ILLLFKNIVPHFTFFVRYTYKELSFQTYRDVHVILVDRSSYKVPVDEAGYINLNGQ 172
 QY 177 CFNSRL-NTGGNLLTYVDRDDINMTVFLQPDVGVTPDKRVGSQPELYLEPGWFGSYR 235
 DB 173 CFSAAIRNOGINRVYHKDDNTNMRLYLXELFGSTINTRYISTPDTQFTYTHW-LYK 231
 QY 236 RRTTVNCELMDMFARSNPPDFVATGDTVEMSPWGSDEDDHENKMKHPWFSVNNY 295
 DB 232 SSSSINCIVDTMGKSDYPYDNFILTGESVEISPPFNCTS--KEVNEQMYFMSKSNY 289
 QY 296 KV---VDYQN--RGTVPLGKTRIFLDREBYTLSWEKHLKMSYCPLTLKAFYNGIQTEH 350
 DB 290 TMLEKLDPEKPKTIP--TIAFLQKGDTLFSWEVKQNTNCKYTAWTKKHALRADM 346
 QY 351 SGSYHFWANDITASFTTSKEDMKFNFTYHCLNEEIKAEIEKKY-AKVNSTHSGYDGLKY 409
 DB 347 TNSYHFMKMDTATWTTKNTINLTGTEYECVKNDIEKYITDTFQSKYNNTHKNENY 406
 QY 410 FKTGGLYLWQPLIQNRLLDANKLNNETYRSRRRQAEISTTDPMMWMTGNGAGGEYS 469
 DB 407 YETTTDLILWQPIVRKSIKELKEFINETQOHTRYKR-----EIDG 447
 QY 470 ENSITVAQVOYADNLRIRINNILEDLSKAWCREQRAALVWNLKINPTSVMNINR 529
 DB 448 SESLVYASLQYMDALREYNAGFAQIAEAWCEDQKRTNEVLSELAKINPSNVMYIDK 507
 QY 530 PVSAREIGDIVISNCIVVDQTSVLSHKSRL--SASDEKCFSPPTVFKPMNDSTI- 585
 DB 508 SUSAKLVGDAISVSSCVNVQSVKHKDMRIYANGTANRETCSRPVTFESSNNSVQ 567
 QY 586 -YKGLGVNNEILLTLYLETQENTYFYFOAKTDMYIYKNYEHKLTVPVLSITLDTFI 644
 DB 568 QFTGQLGPRNEILLGTHRVKCEKERNVSVKVPFAGKEAYFFNYTYTKTVNISDINNVDTFI 627
 QY 645 ALNFTLLENVDFKVIELYTRDEKRLSNVFDIETMFREYNYAQRVSGRLKDLDLSTNRN 704
 DB 628 HLNKPLENTDFEVLRMYSKNELAQANIFDLESLLIDINSYRSALYNIES---RIAPRK 684

Qy	705	QFVDPAGSLMDLGLGAVGTQVNVASGVAT-LFSSIVTGTGFINFIKNPFGG---MLMIIVVI	760
Db	685	DYVSGVDSPFHALGICAPGLGAALGMATGAVTDLTGIFSGFNKPFGLFSLFFVLLV	744
Qy	761	GVLFALYFLTKTKIYETAPIKMIYP-----EIDKLKERECK-----	797
Db	745	FLIFSYY--RKNYI-TNPVALFPYANSSGVVISNTHSYVEETNNKQSENDKPKDTS	801
Qy	798	---SEIAPISEBELERIVLAMHIHQONSHMETKTKDKPDSILTRAQNM	843
Db	802	NAVSEGSANKYSQEDAVCMLMAIKNLGDAYRRKNTTKPSPVLDKIRHL	850

RESULT	44
Q9WT25	
ID	Q9WT25
AC	PRELIMINARY; PRT; 830 AA.
DT	Q9WT25;
DT	01-NOV-1999 (TrEMBLrel. 12, Created)
DT	01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT	01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
GN	Glycoprotein B.
GN	U39.
OS	Human herpesvirus 6.
OC	Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC	Betaherpesvirinae; Roseolovirus.
OX	NCBI_taxid=10368;
RN	[1]
RP	SEQUENCE FROM N.A.
RC	STRAIN=HST;
RC	Isegawa Y.;
RL	Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.
RN	[2]
RP	SEQUENCE FROM N.A.
RC	STRAIN=HST;
RC	MEDLINE-994123119; PubMed=10482554;
RL	Isegawa Y., Mukai T., Nakano K., Kagawa M., Chen J., Mori Y.,
RA	Sunagawa T., Kawanishi K., Sashihara J., Hata A., Zou P., Kosuge H.,
RA	Yamanishi K.;
RA	"Comparison of the complete DNA sequences of human herpesvirus 6
RL	variants A and B.,"
RL	J. Virol. 73:8053-8063(1999).
DR	EMBL; AB021506; BAA78260.1; -.
DR	InterPro; IPR000234; Glycoprot_B.
DR	Pfam; PF00806; Glycoprotein B; 1.
DR	ProDom; PD000693; Glycoprot B; 1.
SQ	SEQUENCE 830 AA; 93136 MW; 608CCF3FDFC31578 CRC64;

Query Match	23.1%;	Score	1052.5;	DB 12;	Length	830;			
Best Local Similarity	32.4%;	Pred.	No. 8.1e-58;						
Matches	245;	Conservative	157;	Mismatches	295;	Indels	59;	Gaps	15
Qy	83	FPRVCSAGVGCDVRFQTDHVC-PDASDMVHSEGLLIYKONIIPMFVRVKRVKVVTT	141						
	:	: :: :	:	:	:	:	:	:	:
Db	36	YPFRICSTAKGDLMPDRDISCPYSNAKMSEGFFIIYKTNIETVPFVRTYKNELTF	95						
	:	:	:	:	:	:	:	:	:
Qy	142	STWANGIYSDITNOHTFKYIEPWETEKMDTIIQCENSLRLN--TGGNLLTVVDRDDIN	199						
	:	:	:	:	:	:	:	:	:
Db	96	QTSRDVGYYFLDRTVWGAMPVEANLNSRAQCISAVAIKRPDGTVESAHEDNNKN	155						
	:	:	:	:	:	:	:	:	:
Qy	200	MTVLQPVDGVTPDKRYGSOPELYLEPGFWGSYRRRTTVNCELMDMFAARNPPDFVF	259						
	:	:	:	:	:	:	:	:	:
Db	156	ETLEFPPLNFKSVTNKRFTTTKEPYFARGPLW-LYSTSTSLNCIVTEATAKAPYSFYA	214						
	:	:	:	:	:	:	:	:	:
Qy	260	TATGDTVMSPFWSGED-DHENKMHKEPWFVSVINNYKV-DYON--RGTVPLGKTRIEL	315						
	:	:	:	:	:	:	:	:	:
Db	215	LTTGEIVGESPPFDGNGKHFAEPLEK---LTILENTMIEDLMGNMGATTLVRIAEFL	271						
	:	:	:	:	:	:	:	:	:
Qy	316	DREEVTLSEKHLKNMSYCPTLWKAFYNGIQTEHSGSYHFVANDITASFTTSKEDMKEF	375						
	:	:	:	:	:	:	:	:	:
Db	272	EKGDTLPSEWEIKENESVCMUKHWTTVTGHURATDETYHFIKELTAAPVTPROSLINT	331						
	:	:	:	:	:	:	:	:	:
Qy	376	NTTYHCLNBEIKAEIEKKY-AKVNSTHSKYGDLYKFXTDGGLYLVWLOPLIONELLDAKNK	434						

[illegible]

	Query Match	23.0%	Score 1047.5;	DB 12;	Length 921;
	Best Local Similarity	28.6%;	Pred. No. 1.9e-57;		
	Matches 273;	Conservative 186;	Mismatches 353;	Indels 143;	Gaps 25;
Qy	14 LWLYQVALYSLIAET----	GYTSPNTATWSTESPLTGCHYGTHDSSHGCRGNNE	RDS	69	
	: :	:	:	:	:
Dd	1 MWAKIMCLVVCVLLRTPSSGGVGAPSSSTGQPPTPVP	CGSSS-----GRTRRTTT	56		
	: :	:	:	:	:
Qy	70 EQON-----	KNIYC-GPSTFPYRVCSASGVGDVFRFOTDHVC---	PDAS	109	
	: :	:	:	:	:
Dd	57 QQUESTNGTEVIENFTLRLPDVVIGVNISKYPYRVCSMAOQTDLIRFDNRHICTSVKPTKE	116			
	: :	:	:	:	:
Qy	110 DMVHSEGIILLYKONIIPNFRVRKYRKVTVTSTTVNGNYISDSITNOHTFYFKSIPEMETE	169			
	: :	:	:	:	:
Dd	117 DL--DEGMVWVKRNIOAHTFKYRVFOKVLTFRRSYAYIYSTYLLGSNTVEYVPPPMMEVH	174			
	: :	:	:	:	:

